



## **SANDIA REPORT**

SAND2001-1669  
Unlimited Release  
Printed June 2001

# **Moving Least-Squares: A Numerical Differentiation Method for Irregularly Spaced Calculation Points**

Albert Gossler

Prepared by  
Sandia National Laboratories  
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,  
a Lockheed Martin Company, for the United States Department of  
Energy under Contract DE-AC04-94AL85000.

Approved for public release; further dissemination unlimited.



**Sandia National Laboratories**

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

**NOTICE:** This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from  
U.S. Department of Energy  
Office of Scientific and Technical Information  
P.O. Box 62  
Oak Ridge, TN 37831

Telephone: (865)576-8401  
Facsimile: (865)576-5728  
E-Mail: [reports@adonis.osti.gov](mailto:reports@adonis.osti.gov)  
Online ordering: <http://www.doe.gov/bridge>

Available to the public from  
U.S. Department of Commerce  
National Technical Information Service  
5285 Port Royal Rd  
Springfield, VA 22161

Telephone: (800)553-6847  
Facsimile: (703)605-6900  
E-Mail: [orders@ntis.fedworld.gov](mailto:orders@ntis.fedworld.gov)  
Online order: <http://www.ntis.gov/ordering.htm>



SAND 2001-1669  
Unlimited Release  
Printed June 2001

**Moving Least-Squares:  
A Numerical Differentiation Method  
For Irregularly Spaced Calculation Points**

Albert Gossler  
Thermal/Fluid Computational Engineering Sciences Department  
Sandia National Laboratories  
P.O. Box 5800  
Albuquerque, New Mexico 87185-0835

**ABSTRACT**

Numerical methods may require derivatives of functions whose values are known only on irregularly spaced calculation points. This document presents and quantifies the performance of Moving Least-Squares (MLS), a method of derivative evaluation on irregularly spaced points that has a number of inherent advantages. The user selects both the spatial dimension of the problem and order of the highest conserved moment. The accuracy of calculations is maintained on highly irregularly spaced points. Not required are creation of additional calculation points or interpolation of the calculation points onto a regular grid. Implementation of the method requires the use of only a relatively small number of calculation points. The method is fast, robust and provides smooth results even as the order of the derivative increases.

This page intentionally left blank.

# CONTENTS

<b>1 INTRODUCTION .....</b>	<b>7</b>
<b>2 FORMULATION .....</b>	<b>9</b>
<b>3 RESULTS .....</b>	<b>15</b>
<b>3.1 Test Parameters .....</b>	<b>15</b>
<b>3.2 Application Results .....</b>	<b>21</b>
<b>3.2.1 Number of Nearest Neighbors <math>N</math> .....</b>	<b>21</b>
<b>3.2.2 Resolution <math>N_s</math> .....</b>	<b>30</b>
<b>3.2.3 Scaled Isolation Distance <math>D</math> .....</b>	<b>40</b>
<b>3.2.4 Gaussian Radius <math>\delta</math> .....</b>	<b>48</b>
<b>3.2.5 Execution Time .....</b>	<b>57</b>
<b>4 SUMMARY .....</b>	<b>61</b>
<b>REFERENCES .....</b>	<b>65</b>

This page intentionally left blank.

# 1 INTRODUCTION

The origin of least-squares methods can be traced back to the beginning of statistical methods—perhaps to the first time a mean was calculated on a set of data. The notion of *moving* least-squares (MLS), analogous to a *moving* average, in which a defined subset of data is used for each data reducing calculation, also has an extensive history. The use of MLS in one-dimensional digital filters, for example, is well established, cf. [1][2][3]. Shepard [4] first applied the method of MLS to the generation of two-dimensional interpolants in a limited low-order case and Lancaster and Salkaukas [5] give a general higher-order approach of the method. Armentano and Duran [6] have recently given error estimates for one-dimensional MLS approximations of functions and their 1<sup>st</sup> and 2<sup>nd</sup> order derivatives. As an alternative to finite-element methods, Nayroles et al. [7] and Belytschko et al. [8] adapted MLS approaches to meshless Galerkin approximations; a general overview of meshless methods including MLS is given by Belytschko, et al. [9]. Variations of MLS within and across disciplines are referred to by different names; in meshless methods names include the Diffuse Element Method (DEM) [7], the Element-Free Galerkin (EFG) method [8] and Partition of Unity Method (PUM) [10][11]; in data filter methods, names include Savitzky-Golay filters [1], and Digital Smoothing Polynomials (DISPO) [3].

The present work is motivated by the development of a three-dimensional, vorticity-based, Lagrangian approach to fluid dynamics, a field to which MLS has been introduced relatively recently by Marshall and Grant [12]. Neither analytical error bounds for the MLS approximation of derivatives of multi-dimensional equations nor extensive analysis of the behavior of MLS estimations in such applications exist in the literature. This paper presents an MLS formulation for multi-dimensional applications and presents a detailed investigation of an MLS three-dimensional application including errors in the approximation of 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> order derivatives.

The governing equation for this fluid dynamics approach for a problem of incompressible, three-dimensional fluid flow of variable density and constant viscosity occupying a region  $V$  is given by the vorticity transport equation

$$\frac{d\vec{\omega}}{dt} \equiv \frac{\partial \vec{\omega}}{\partial t} + (\vec{u} \cdot \vec{\nabla})\vec{\omega} = (\vec{\omega} \cdot \vec{\nabla})\vec{u} - \vec{\nabla} \left( \frac{1}{\rho} \right) \times \vec{\nabla} p + \nu \nabla^2 \vec{\omega}, \quad (1)$$

where the symbol  $\vec{\nabla}$  refers to the gradient and the symbol  $\nabla^2$  refers to the 2<sup>nd</sup> order gradient, the Laplacian. In this approach, calculation points that need not be connected by a mesh represent function values in  $V$ . These calculation points are advected at the local fluid velocity and, as such, are irregularly spaced. [The variables in equation (1) are not of primary concern in this discussion and comprise: vorticity ( $\vec{\omega}$ ), time ( $t$ ), velocity ( $\vec{u}$ ), density ( $\rho$ ), pressure ( $p$ ), and kinematic viscosity ( $\nu$ ).] Numerical evaluation of (1) clearly requires that the gradient and Laplacian be evaluated within  $V$  on the irregularly distributed calculation points.

A weighted residual method for an  $r^{\text{th}}$  order problem  $\vec{y}^r = f(\vec{y}^{r-1}, \vec{y}^{r-2}, \dots, \vec{y}, \vec{x})$  can be arranged in the form

$$\vec{E} = \vec{y}^r - f(\vec{y}^{r-1}, \vec{y}^{r-2}, \dots, \vec{y}, \vec{x}). \quad (2)$$

where  $\vec{E}$  is an “error” or residual vector of the original equation,  $\vec{x}$  is a vector of independent variables and  $\vec{y}^\xi$  are vectors of the  $\xi^{\text{th}}$  derivative of the dependent variable, cf. [13]. This residual can be minimized using any of several methods that cause a weighted average of the residual to vanish. Such approaches vary only in the manner in which the residual is weighted and include: (i) collocation, in which the actual values of the residual at selected points are made equal to zero; (ii) Galerkin’s method, in which the integrals of the residual weighted by selected shape functions are set to zero; (iii) a least-squares approach, in which the integral of the square of the residual is minimized. It cannot be said *a-priori* which approach yields the most accurate solution for a given case. Hence, selection of the approach rests primarily on foreknowledge of the particular application or on other considerations. For example, the collocation approach does not require integration, providing for an easy implementation. However, there is no guarantee that this approach yields a solution that is sufficiently smooth for calculations of derivatives, particularly higher order derivatives. The shape functions in Galerkin’s method can be used to create a symmetric coefficient matrix, often of advantage in finite element methods. In the present application, the MLS approach is selected because it maintains good accuracy on irregularly spaced calculation points [12] and because of its inherent smoothing properties.



## 2 FORMULATION

The use of a Lagrangian numerical approach to solve a transport equation, such as the vorticity transport equation (1), yields values at points  $\bar{\mathbf{x}}_n$ ,  $n = 1, \dots, \bar{N}$ , which, in general, are irregularly spaced. An MLS approach is used here to evaluate the derivatives of a function, say  $f(\bar{\mathbf{x}}, t)$ , at a calculation point  $m$  located at  $\bar{\mathbf{x}}_m$ . In this method, the values  $f_n$  of  $f(\bar{\mathbf{x}}, t)$  on these calculation points are interpolated locally by a polynomial in the components of the position difference  $\bar{\mathbf{x}} - \bar{\mathbf{x}}_m$ .

$$q_m(\bar{\mathbf{x}}, t) = f_m + \sum_{i=1}^k C_{m,i} B_{m,i}(\bar{\mathbf{x}} - \bar{\mathbf{x}}_m). \quad (3)$$

In (3), the index  $m$  denotes the point about which the interpolation is performed,  $C_{m,i}$  denotes a set of  $k$  undetermined coefficients of the polynomial, and  $B_{m,i}(\bar{\mathbf{x}} - \bar{\mathbf{x}}_m)$  are the associated basis functions. The value of  $k$  is the total number of combinations with repeat  $K^R$  [14] possible for the higher order terms in dimensions  $d$  and order  $h$ , given by

$$k = \sum_{j=1}^h K^R(d, j) = \sum_{j=1}^h \frac{(d-1+j)!}{(d-1)! j!}. \quad (4)$$

The prescribed order  $h$  polynomial fit is equal to the highest moment conserved by the basis functions  $B_{m,i}(\bar{\mathbf{x}} - \bar{\mathbf{x}}_m)$ . The MLS approach using a polynomial fit of prescribed order  $h$  (calculated with a sufficient number of calculation points) represents an  $h^{\text{th}}$  order function exactly and is referred to as an  $h^{\text{th}}$  order MLS fit. The  $k$  basis functions are generated using the relationship

$$B_{m,i} = \frac{(x - x_m)^a (y - y_m)^b (z - z_m)^c}{R_m^{a+b+c}} \quad (5)$$

where  $i = 1, \dots, k$ , and  $a, b, c$  are whole numbers with  $a + b + c = j$ ,  $j = 1, \dots, h$ . For the three-dimensional, second order case  $k|_{d=3, h=2} = 9$  and the associated basis functions are

$$B_{m,1} = \frac{x - x_m}{R_m}, \quad (6a)$$

$$B_{m,2} = \frac{y - y_m}{R_m}, \quad (6b)$$

$$B_{m,3} = \frac{z - z_m}{R_m}, \quad (6c)$$

$$B_{m,4} = \left( \frac{x - x_m}{R_m} \right) \left( \frac{y - y_m}{R_m} \right), \quad (6d)$$

$$B_{m,5} = \left( \frac{x - x_m}{R_m} \right) \left( \frac{z - z_m}{R_m} \right), \quad (6e)$$

$$B_{m,6} = \left( \frac{y - y_m}{R_m} \right) \left( \frac{z - z_m}{R_m} \right), \quad (6f)$$

$$B_{m,7} = \left( \frac{x - x_m}{R_m} \right)^2, \quad (6g)$$

$$B_{m,8} = \left( \frac{y - y_m}{R_m} \right)^2, \quad (6h)$$

$$B_{m,9} = \left( \frac{z - z_m}{R_m} \right)^2. \quad (6i)$$

The parameter  $R_m$  is a length scale associated with the calculation point  $m$ , and can be considered an effective point “radius”. The value of  $R_m$  can be set for each point  $m$  as a function of the local average calculation point spacing or can simply be based on typical spacing between calculation points. Either approach is used to ensure that for small difference components the basis function values do not approach the computer’s floating-point precision (typically  $10^{-6}$  for single precision or  $10^{-12}$  for double precision).

Implementing the MLS method in higher or lower dimensions or other orders is straightforward. For example, to attain third-order polynomial fit ( $h=3$ ) in three dimensions ( $d=3$ ) requires 19 higher order terms ( $k|_{d=3, h=3}=19$ ). The corresponding basis functions are given by those in equations (6) with additional functions given by

$$B_{m,10} = \left( \frac{x - x_m}{R_m} \right)^2 \left( \frac{y - y_m}{R_m} \right), \quad (7a)$$

$$B_{m,11} = \left( \frac{x - x_m}{R_m} \right) \left( \frac{y - y_m}{R_m} \right)^2, \quad (7b)$$

$$B_{m,12} = \left( \frac{x - x_m}{R_m} \right)^2 \left( \frac{z - z_m}{R_m} \right), \quad (7c)$$

$$B_{m,13} = \left( \frac{x - x_m}{R_m} \right) \left( \frac{z - z_m}{R_m} \right)^2, \quad (7d)$$

$$B_{m,14} = \left( \frac{y - y_m}{R_m} \right)^2 \left( \frac{z - z_m}{R_m} \right), \quad (7e)$$

$$B_{m,15} = \left( \frac{y - y_m}{R_m} \right) \left( \frac{z - z_m}{R_m} \right)^2, \quad (7f)$$

$$B_{m,16} = \left( \frac{x - x_m}{R_m} \right) \left( \frac{y - y_m}{R_m} \right) \left( \frac{z - z_m}{R_m} \right), \quad (7g)$$

$$B_{m,17} = \left( \frac{x - x_m}{R_m} \right)^3, \quad (7h)$$

$$B_{m,18} = \left( \frac{y - y_m}{R_m} \right)^3, \quad (7i)$$

$$B_{m,19} = \left( \frac{z - z_m}{R_m} \right)^3, \quad (7j)$$

The coefficients  $C_{m,i}$  of the polynomial (3) are obtained by a localized least-squares procedure, in which the “error”  $E_m$  is expressed as

$$E_m \equiv \sum_{n=1}^N L_{mn} [f_n - q_m(\bar{\mathbf{x}}_n, t)]^2. \quad (8)$$

where  $N$  is the number of calculation points, the “nearest neighbors”, about point  $\bar{\mathbf{x}}_m$  used in the MLS fit. The “localization parameter”  $L_{mn}$  weights the contribution of different points to

the error  $E_m$ . The value of  $L_{mn}$  can be set equal to unity for the  $N$  nearest neighbors of  $m$  and zero elsewhere or its value may be set to decay with distance from  $\bar{\mathbf{x}}_m$  using any convenient function.

Minimization of  $E_m$  with respect to each of the coefficients  $C_{m,i}$  yields a  $k \times k$  system of linear equations of the form

$$\sum_{j=1}^k G_{m,ij} C_{m,j} = U_{m,i}, \quad i = 1, \dots, k, \quad (9)$$

where

$$G_{m,ij} = \sum_{n=1}^N L_{nm} B_{nm,i} B_{nm,j}, \quad (10)$$

$$U_{m,i} = \sum_{n=1}^N (f_n - f_m) L_{nm} B_{nm,i} \quad (11)$$

$$B_{nm,i} \equiv B_{m,i}(\bar{\mathbf{x}}_n - \bar{\mathbf{x}}_m). \quad (12)$$

Solution of the system (9) yields

$$C_{m,i} = \sum_{j=1}^k G_{m,ij}^{-1} U_{m,j}, \quad (13)$$

where  $G_{m,ij}^{-1}$  is the inverse of  $G_{m,ij}$ . Upon solving for coefficients  $C_{m,i}$ , the derivatives of  $f_m$  are approximated by differentiating the polynomial fit (3). As examples, for the three-dimensional, second order polynomial case, the first order derivatives are given by

$$\left. \frac{\partial f_n(\bar{\mathbf{x}})}{\partial x} \right|_{\bar{\mathbf{x}}_n = \bar{\mathbf{x}}_m} = \frac{C_{m,1}}{R_m}, \quad (14a)$$

$$\left. \frac{\partial f_n(\bar{\mathbf{x}})}{\partial y} \right|_{\bar{\mathbf{x}}_n = \bar{\mathbf{x}}_m} = \frac{C_{m,2}}{R_m}, \quad (14b)$$

$$\left. \frac{\partial f_n(\bar{\mathbf{x}})}{\partial z} \right|_{\bar{\mathbf{x}}_n = \bar{\mathbf{x}}_m} = \frac{C_{m,3}}{R_m}, \quad (14c)$$

and the second order derivatives are given by

$$\left. \frac{\partial^2 f_n(\vec{x})}{\partial x^2} \right|_{\vec{x}_n = \vec{x}_m} = \frac{2C_{m,7}}{R_m^2}, \quad (15a)$$

$$\left. \frac{\partial^2 f_n(\vec{x})}{\partial y^2} \right|_{\vec{x}_n = \vec{x}_m} = \frac{2C_{m,8}}{R_m^2}, \quad (15b)$$

$$\left. \frac{\partial^2 f_n(\vec{x})}{\partial z^2} \right|_{\vec{x}_n = \vec{x}_m} = \frac{2C_{m,9}}{R_m^2}. \quad (15c)$$

For the three-dimensional, third order polynomial case, the first and second order derivatives are again given by equations (14) and (15) and third order derivatives are given by

$$\left. \frac{\partial^3 f_n(\vec{x})}{\partial x^3} \right|_{\vec{x}_n = \vec{x}_m} = \frac{6C_{m,17}}{R_m^3}, \quad (16a)$$

$$\left. \frac{\partial^3 f_n(\vec{x})}{\partial y^3} \right|_{\vec{x}_n = \vec{x}_m} = \frac{6C_{m,18}}{R_m^3}, \quad (16b)$$

$$\left. \frac{\partial^3 f_n(\vec{x})}{\partial z^3} \right|_{\vec{x}_n = \vec{x}_m} = \frac{6C_{m,19}}{R_m^3}. \quad (16c)$$

The coefficient matrix in equation (9) can be solved using a variety of linear equation solvers including Gauss-Jordan elimination, Gaussian elimination, LU decomposition and Cholesky decomposition. However, under typical conditions the condition number of the coefficient matrix may become very large. Any matrix is singular if its condition number is infinite and can be defined as ill-conditioned if the reciprocal of its condition number approaches the computer's floating-point precision. In such cases, the cited solvers may fail or may return a highly inaccurate solution. To avoid such numerical problems, a singular value decomposition (SVD) linear equation solver is sometimes recommended for use in conjunction with the MLS method. The SVD solver identifies equations in the matrix that are, within a specified tolerance, redundant (linear combinations of the remaining equations) and eliminates them, thereby improving the condition number of the matrix. The reader is referred to reference [15], Chapter 15 for a helpful discussion of SVD pertinent to linear least-squares problems. To examine their relative merits, two linear equation solvers are used in this report: Gauss-Jordan elimination (GJE) with full pivoting, a direct solver that is robust and relatively simple; SVD, a solver that remains robust even for initially ill-conditioned sets of linear equations.

This page intentionally left blank.

### 3 RESULTS

This section extensively explores a Moving Least Squares application in order to gain working knowledge of the approach and so develop helpful comments, small hints and general advice to make MLS a practical tool.

#### 3.1 Test Parameters

The MLS approach is tested by using it to evaluate the 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> derivatives of a three-dimensional Gaussian test function  $f$  given by

$$f(x, y, z) = \exp \left( -\frac{(x-0.5)^2 + (y-0.5)^2 + (z-0.5)^2}{\delta^2} \right). \quad (17)$$

The parameter  $\delta$  is the Gaussian “radius” and determines the width of the single “wave” in this smoothly varying function. As  $\delta$  decreases, the function “narrows.” Before introducing a graph of this function, a “thin center slice” will be defined as a nearly one-dimensional function space along the  $x$ -axis within the region  $(0.5 - \Delta x) < y < (0.5 + \Delta x)$  and  $(0.5 - \Delta x) < z < (0.5 + \Delta x)$ , where  $\Delta x$  is the average spacing between calculation points.

Figure 1 presents a graph of the Gaussian function given by equation (17) on the test function space used throughout this section, a  $1 \times 1 \times 1$  cube extending from the origin to 1 on each Cartesian axis. The radius used for this graph is given by  $\delta^2 = 0.035$ . This radius is used for all tests within this section except for tests on the effect of varying the radius. While this radius yields the widest Gaussian function considered, this function itself is not of primary interest; its derivative functions are. The waves of higher order derivatives of this function become increasingly narrow with the ratio of “wave amplitude” to “wave width” becoming increasingly severe. The test function gradient  $\nabla f$  is shown in Figure 2, the Laplacian  $\nabla^2 f$  in Figure 3, and 3<sup>rd</sup> order gradient  $\nabla^3 f$  in Figure 4. The wave amplitude to width ratio  $\Omega$  can be seen to increase rapidly from approximately  $\Omega = 10$  for  $\nabla f$  to a value of approximately  $\Omega = 325$  for  $\nabla^2 f$  and then to a value of approximately  $\Omega = 2400$  for  $\nabla^3 f$ .

Because MLS is an approach applied to solution techniques in which function values are known generally on irregularly spaced calculation points, a topic of interest is the effect of the randomness of the point arrangement on the accuracy and stability of the result. Toward this end, functions are tested using both evenly spaced calculation points with uniform spacing  $\Delta x$  and with randomized calculation points with average local spacing  $\Delta x$ . The randomized point locations are generated by perturbing the uniformly spaced points in each Cartesian direction by  $\alpha \Delta x$  where  $\alpha$  is a random number in the range  $0 \leq \alpha \leq \varepsilon$ .

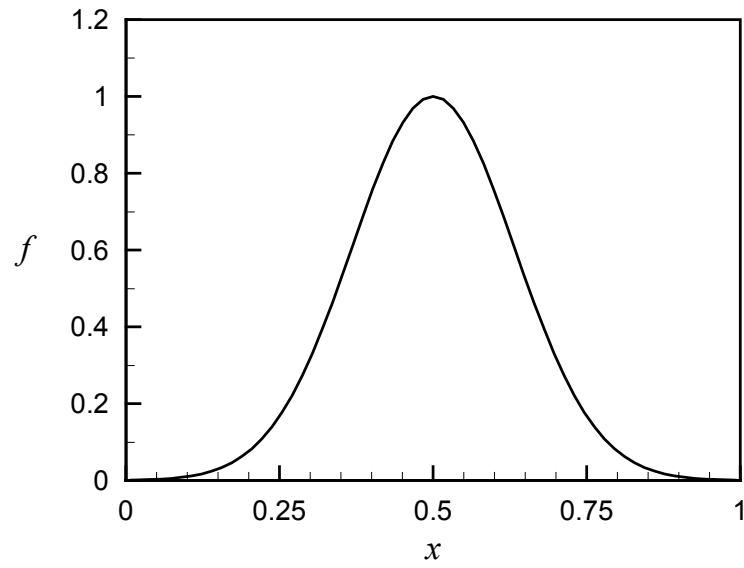


Figure 1. A thin center slice of the Gaussian test function  $f$  with  $\delta^2 = 0.035$ .

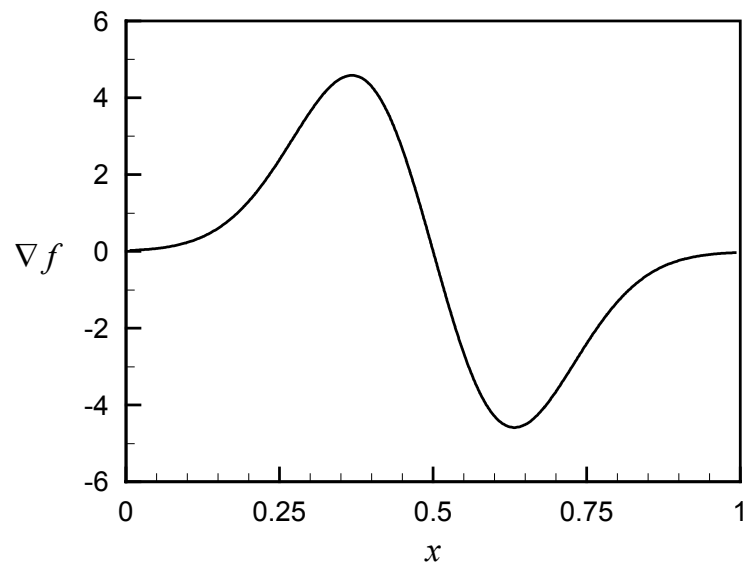


Figure 2. A thin center slice showing the gradient  $\nabla f$  of the Gaussian test function  $f$  with  $\delta^2 = 0.035$ .



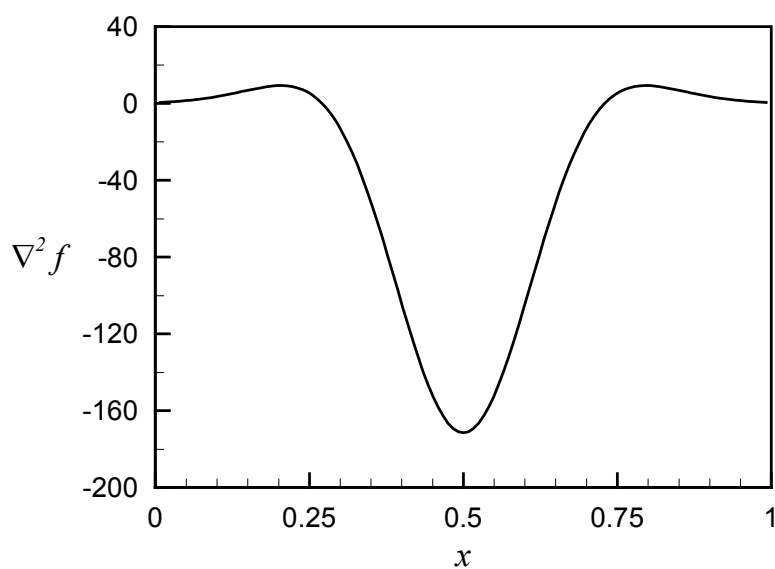


Figure 3. A thin center slice showing the Laplacian  $\nabla^2 f$  of the Gaussian test function  $f$  with  $\delta^2 = 0.035$ .

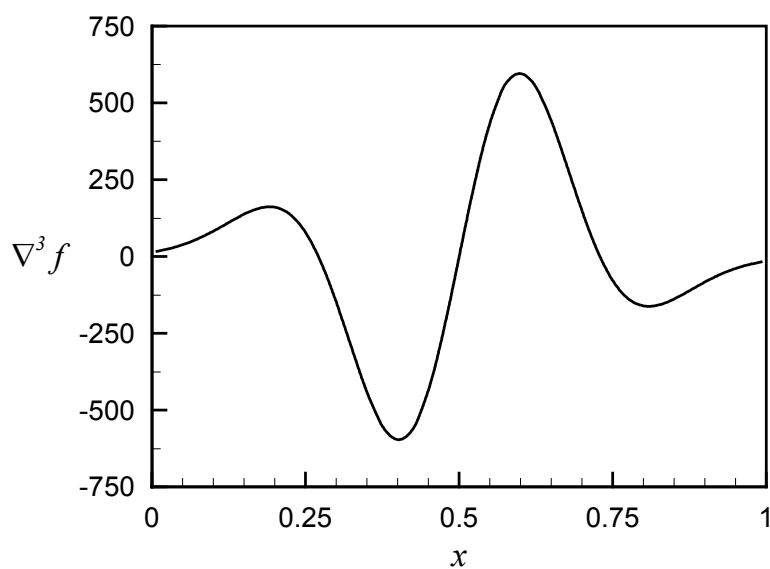


Figure 4. A thin center slice showing the third order gradient  $\nabla^3 f$  of the Gaussian test function  $f$  with  $\delta^2 = 0.035$ .

Results of applying maximum randomization perturbations of  $\varepsilon = 0.05$  and  $\varepsilon = 0.50$  on approximately 1000 point positions are shown in Figure 5. The randomization with  $\varepsilon = 0.05$  is barely detectable by eye whereas the randomization with  $\varepsilon = 0.50$  is plainly evident. Randomization with  $\varepsilon = 0.50$  represents the maximum randomization possible such that the average spacing local to any calculation point remains equal to the uniform spacing  $\Delta x$ .

The effective point radius  $R_m$  in equation (5) is set equal to  $\Delta x$ . While the effect of changing  $R_m$  is not explicitly explored, informal observation shows varying this parameter within an order of magnitude greater or less than  $\Delta x$  has virtually no effect on the MLS results presented.

The value of  $L_{mn}$  in equation (8) is set equal to unity for the  $N$  nearest neighbors of  $m$  and zero elsewhere. The effect of this parameter is a complex consideration. In general, one-dimensional MLS fits, as are used in digital signal processing, generate smoother function approximations  $f_n$  (by removing high frequency variations in the sample function) for  $L_{mn}$  functions that decay with distance from  $\bar{\mathbf{x}}_m$ . The interested reader is referred to reference [3], Chapter 5, for an introductory discussion of this topic.

The parameters  $N_s$  and  $\bar{N}$  are used in investigating resolution effects later in this section. The number of points per side of each face of the  $1 \times 1 \times 1$  test space is given by  $N_s$ . The total number of points  $\bar{N} = N_s^3$  resolving the three-dimensional functions is set by varying  $N_s$  within the range  $21 \leq N_s \leq 51$  such that the range of  $\bar{N}$  spans nearly 1.5 magnitudes,  $9261 \leq \bar{N} \leq 132651$ .

A “scaled isolation distance”  $D$  used in later tests is introduced here. The distance between an isolated calculation point and its nearest neighbor is measured as the scaled isolation distance  $D$  times the average calculation point spacing  $\Delta x$ .

Two parameters are calculated for comparison of test results. One parameter is  $C_{max}$ , the maximum condition number of the coefficient matrix  $G_{m,ij}$  of equation (9) prior to elimination of redundant and nearly redundant equations using SVD. The other parameter is an  $\ell_2$  norm error, a scaled rms (root-mean-square) calculation given by

$$E_{rms} = \frac{1}{\max[g(\bar{\mathbf{x}})]} \sqrt{\frac{1}{\bar{N}} \sum_{i=1}^{\bar{N}} [g'(\bar{\mathbf{x}}_i) - g(\bar{\mathbf{x}}_i)]^2} . \quad (18)$$

where  $g$  and  $g'$  represent the analytic and calculated values, respectively, of  $\nabla f$ ,  $\nabla^2 f$ , or  $\nabla^3 f$ .

For all tests, floating-point calculations are performed using double precision. The performance of MLS using either the GJE or SVD linear equation solver is quantified by evaluating graphs of calculation point values and by evaluating the error and the condition number, summarized in Table 1, as a function of the independent parameters summarized in Table 2.

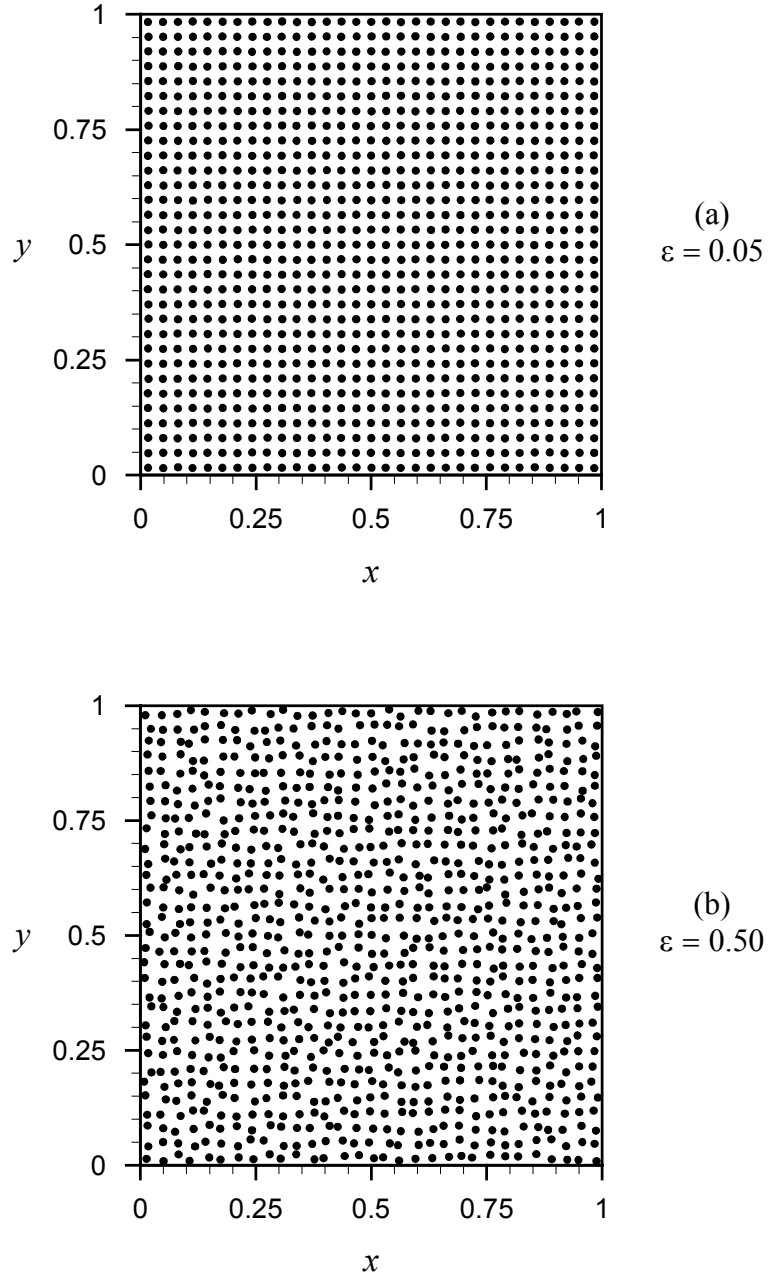


Figure 5. Approximately 1000 point positions randomized using different maximum perturbation amplitudes  $\varepsilon$ .

	Description
$C_{max}$	Maximum condition number.
$E_{rms}$	Scaled rms calculation given by equation (18).

Table 1. Values evaluated in the MLS tests.

	Description	Range
$N$	Number of nearest neighbors used in the MLS fit.	$1 < N < 100$
$N_s$	Number of calculation points per side of each face of the test volume. The test volume is resolved by the total number of calculation points $\bar{N} = N_s^3$ .	$21 \leq N_s \leq 51$
$\varepsilon$	Maximum randomization perturbation. Calculation points of uniform spacing $\Delta x$ are perturbed in each Cartesian direction by $\alpha \Delta x$ where $\alpha$ is a random number in the range $0 \leq \alpha \leq \varepsilon$ .	$0 \leq \varepsilon \leq 0.50$
$D$	Scaled isolation distance: the distance between an isolated calculation point and its nearest neighbor is measured as $D$ times the average calculation point spacing $\Delta x$ .	$1 \leq D \leq 10$
$\delta$	The radius of the Gaussian test function $f$ given by equation (17).	$0.005 \leq \delta^2 \leq 0.035$

Table 2. Summary of parameters varied in MLS tests.

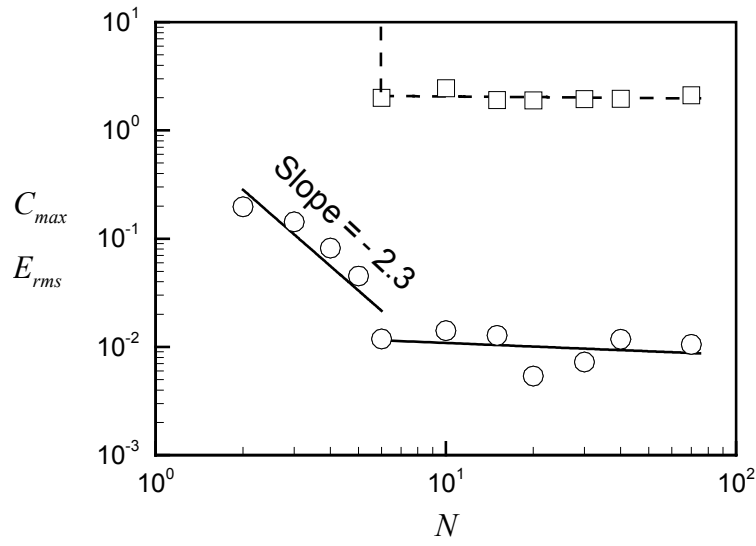
## 3.2 Application Results

The titles of the subsections within this section indicate the parameter that is the focus of tests within that subsection. This primary parameter is varied and results presented and discussed. Also, generally the two different linear equation solvers are used and supplementary parameters such as the maximum randomization perturbation  $\varepsilon$  and the MLS order of fit  $h$  are varied to help investigate the effects of the primary parameter of a subsection. Use of the GJE solver is explicitly stated; otherwise, use of the SVD solver is implied. A short summary of main results appears at the end of each subsection.

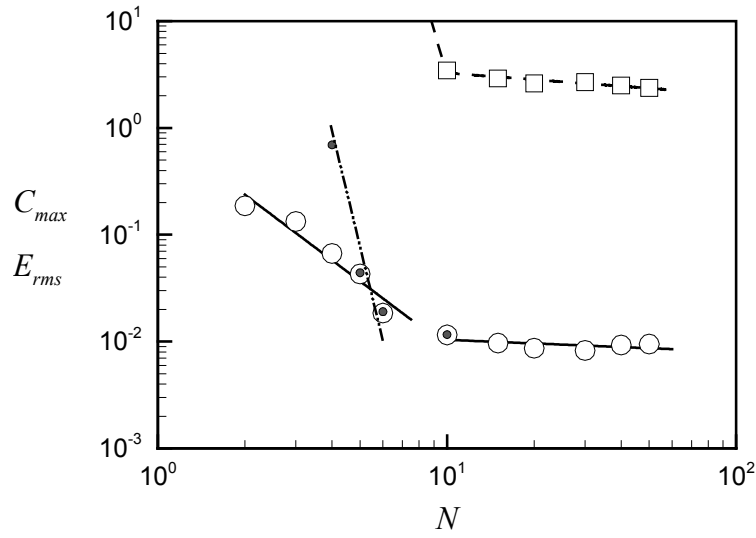
### 3.2.1 Number of Nearest Neighbors $N$ .

The first (and last) consideration in this application is choosing the number of nearest neighbors  $N$  of equation (8) to use in the MLS fit. For this series of tests, the resolution is set to  $N_s = 51$ . Beginning with a 1<sup>st</sup> order MLS fit and using either the GJE solver or the SVD solver, the condition number  $C_{max}$  and the error  $E_{rms}$  are calculated as a function of the number of nearest neighbors  $N$  for uniformly spaced calculation points ( $\varepsilon = 0$ ) and for the randomized points with  $\varepsilon = 0.50$ . Figure 6 presents the results for these tests using the uniformly spaced calculation point and randomized points, respectively. Randomized points are seen to “smooth out” the coefficient condition numbers  $C_{max}$ ; with decreasing  $N$ ,  $C_{max}$  increases abruptly for uniformly spaced points and increases gradually for randomized points. The number of nearest neighbors  $N$  required to achieve the lowest possible error  $E_{rms}$  is slightly greater for the randomized points. The results using the GJE solver are indistinguishable from those using the SVD solver presented in Figure 6(a) and 6(b) in the region where the condition number  $C_{max}$  is bounded to values  $C_{max} \ll 10^{12}$ . For uniformly spaced calculation points, the GJE solver abruptly fails (for  $N \leq 5$ ) due to the matrix becoming effectively and abruptly singular. For the randomized points, as the condition number  $C_{max}$  quickly increases, the GJE solver generates large  $E_{rms}$  errors (for  $4 \leq N \leq 5$ ) and then fails (for  $N \leq 3$ ) as  $C_{max}$  becomes unbounded. Selected results using the GJE solver are included in Figure 6(b) to highlight the difference in behavior between the two solvers.

For uniformly spaced and randomized calculation points, the SVD solver does not fail as the condition number  $C_{max}$  increases. Rather, the SVD solver provides solutions with the smallest possible least-square coefficient values (as opposed to values approaching infinity!). The errors  $E_{rms}$  of these solutions increase substantially and gradually with decreasing  $N$ . Lines drawn through the error values  $E_{rms}$  are broken into two parts corresponding to regions in which the condition number is sharply increasing or nearly constant. In the region of sharply increasing condition number  $C_{max}$ , error  $E_{rms}$  increases with decreasing  $N$  at slightly greater than 2<sup>nd</sup> order (as labeled) for the uniformly spaced calculation points and with 2<sup>nd</sup> order (not labeled) for the randomized points.



(a)  
uniformly  
spaced  
points



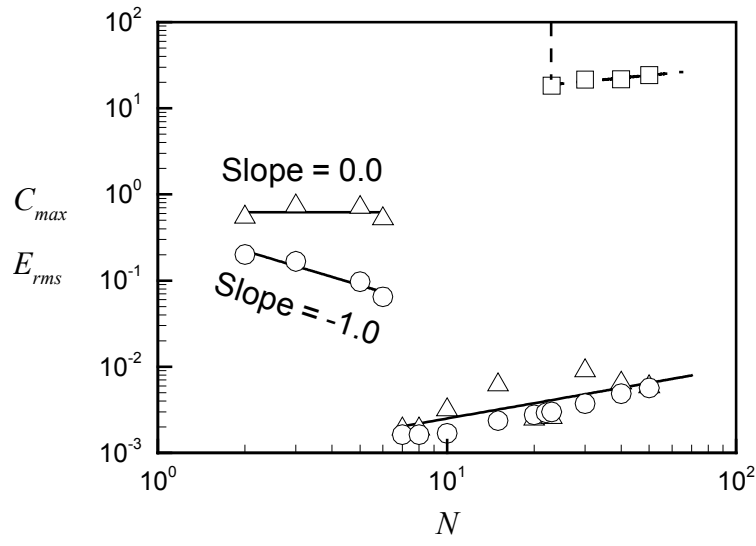
(b)  
randomized  
points  
 $\epsilon = 0.50$

Figure 6. Results for the 1<sup>st</sup> order MLS fit. The condition number  $C_{max}$  (---□---) and the error  $E_{rms}$  for the gradient (SVD:—○— ; GJE:---●---).

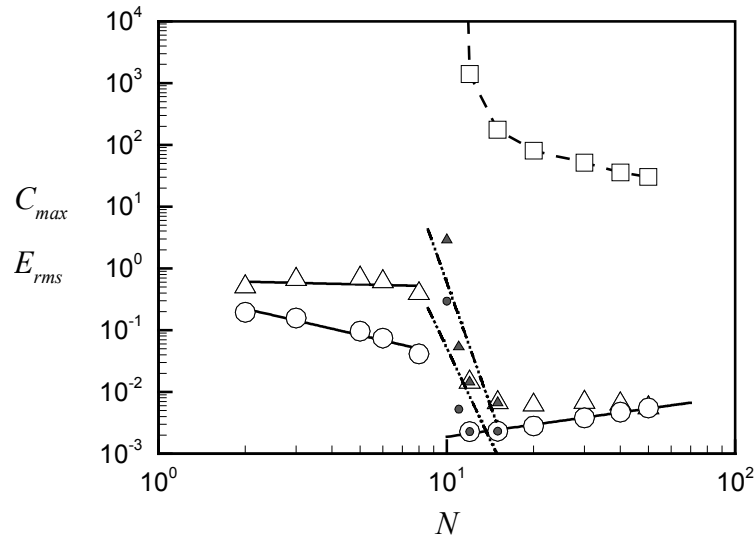
If the basis functions for the  $N$  points were unique, the linear equations in the coefficient matrix would be unambiguous and the condition number  $C_{max}$  and the error  $E_{rms}$  would remain roughly constant. Hence, the basis functions are effectively redundant when evaluated numerically. The randomization of calculation point positions helps to maintain unique contribution of the associated basis functions. This result, while perhaps not surprising, nonetheless communicates an uncommon knowledge: for the MLS approach, and possibly other minimization approaches, calculating function values using randomized points can be more accurate and more stable than using uniformly spaced points.

The number of nearest neighbors to use in the 1<sup>st</sup> order MLS fit can now be chosen based on the presented results. Assuming for the moment that there must be some calculation penalty (to be evaluated later) in using more nearest neighbor calculation points for the fit, a minimum number of points is selected that, in this case, (i) maintains an error of approximately  $E_{rms} = 10^{-2}$  and (ii) avoids solving an ill-conditioned matrix. That number is  $N = 7$ . Using 7 points will provide a sufficient number of unique basis equations to determine the unknown MLS coefficient matrix of equation (9) for this application under all conditions.

For similar conditions, 1<sup>st</sup> and 2<sup>nd</sup> order gradients are calculated using a 2<sup>nd</sup> order MLS fit. Comparing the condition numbers  $C_{max}$  of results for uniformly spaced calculation points with those for randomized points, shown in Figure 7, the tendency of randomization is again to smooth out (lessen the magnitude and make more gradual the change in) the condition numbers  $C_{max}$ . In the region where the condition number  $C_{max}$  is bounded to values  $C_{max} \ll 10^{12}$ , the results using the GJE solver are indistinguishable from those using the SVD solver presented in Figure 7(a) and Figure 7(b). For uniformly spaced calculation points, the GJE solver fails (for  $N \leq 22$ ) as the matrix abruptly becomes singular, consistent with the 1<sup>st</sup> order MLS results. The condition number  $C_{max}$  when using randomized points quickly increases and the GJE solver generates large  $E_{rms}$  errors (for  $N = 10$  and  $N = 11$ ) and then fails (for  $N \leq 9$ ) as  $C_{max}$  increases without bound. In contrast, the SVD solver again provides solutions with increased  $E_{rms}$  errors for poorly conditioned (relatively large condition numbers but  $C_{max} < 10^{12}$ ) and even for ill-conditioned matrices ( $C_{max} > 10^{12}$ ). To demonstrate this change in behavior between results generated using the two solvers, selected results using the GJE solver are included in Figure 7(b).



(a)  
uniformly  
spaced  
points



(b)  
randomized  
points  
 $\epsilon = 0.50$

Figure 7. Results for the 2<sup>nd</sup> order MLS fit. The condition number  $C_{max}$  (---□---) and the error  $E_{rms}$  for the gradient (SVD:—○— ; GJE:---●---) and the Laplacian (SVD:—△— ; GJE:---▲---).



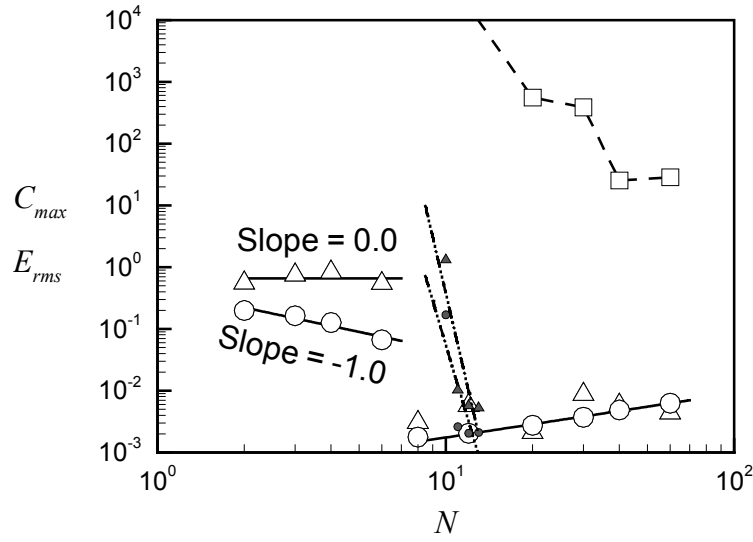


Figure 8. Results using the same parameters as those of Figure 7 except for randomization, which is reduced to  $\varepsilon = 0.05$ . The condition number  $C_{max}$  (---□---) and the error  $E_{rms}$  for the gradient (SVD:—○—; GJE: ---●---) and the Laplacian (SVD:—△—; GJE: ---▲---) for the 2<sup>nd</sup> order MLS fit.

For the 2<sup>nd</sup> order fit and randomized calculation points, if the criterion is simply that error  $E_{rms}$  remains below  $10^{-2}$  for uniform or randomized calculation points, then  $N = 14$  is adequate. Let the acceptable minimum number of nearest neighbors for a 2<sup>nd</sup> order MLS fit using randomized points be  $N_{min} = 14$  for future reference. If the criterion is added to avoid solving an ill-conditioned matrix, a value of  $N = 23$  might be chosen. Let the acceptable maximum number of nearest neighbors for a 2<sup>nd</sup> order MLS fit using randomized points be  $N_{max} = 23$ . Selecting the actual value of  $N$  to be used will be deferred until after review of test results presented in succeeding subsections. Formally applying the minimum and maximum nearest neighbors criteria to the 1<sup>st</sup> order MLS fit discussed above gives  $N_{min} = N_{max} = 7$  for that case.

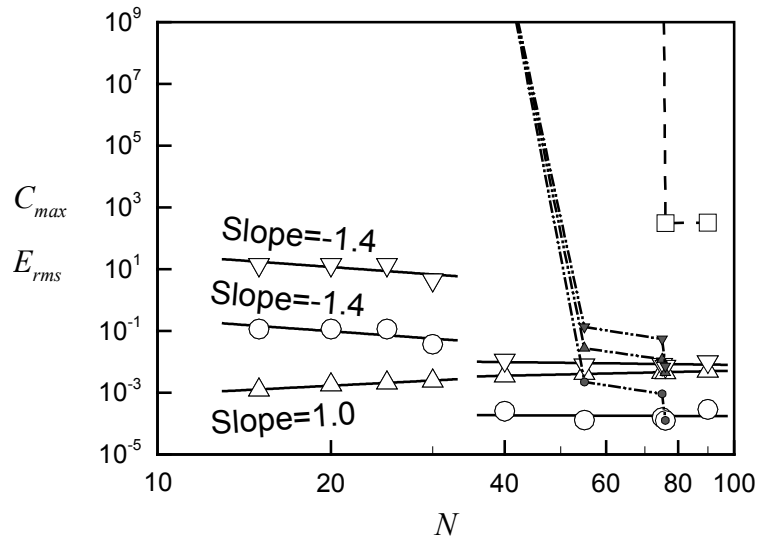
Randomization with  $\varepsilon = 0.50$  shown in Figure 7(b) is reduced to  $\varepsilon = 0.05$  and shown in Figure 8 to gage the effect of the degree of randomization. The trends are similar in these figures: even minimal randomization smoothes the condition number  $C_{max}$ ; error  $E_{rms}$  values are in close agreement between the figures using either the GJE or SVD solvers. Considering the region of greatest interest where the lowest number of nearest neighbors  $N$  provides the lowest error  $E_{rms}$ , a severe reduction in condition number  $C_{max}$  is achieved by randomizing calculations points even slightly while there is only a small and gradual change in the  $E_{rms}$  errors as randomization is increased. As with the greater randomization,  $E_{rms}$  errors when using the GJE solver are indistinguishable from those when using the SVD

solver in the region where the condition number is bounded to values  $C_{max} \ll 10^{12}$ . Included in Figure 8 are selected GJE results to show the divergent behavior of the two solvers.

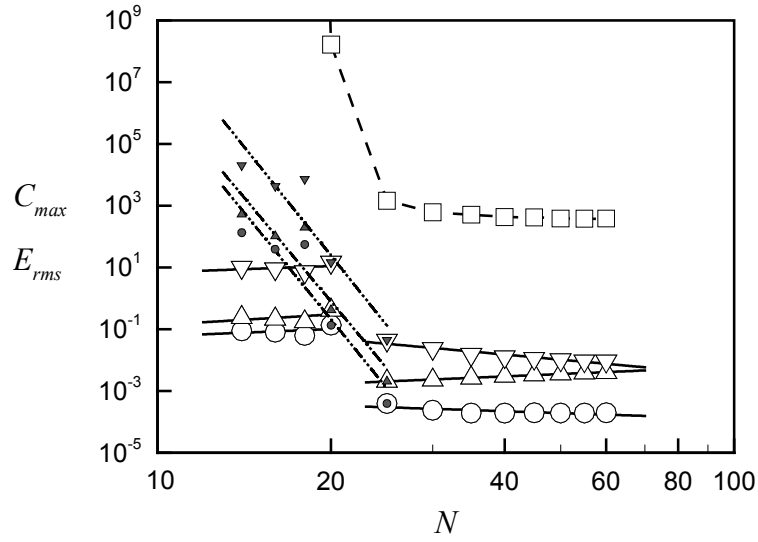
The 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> order gradients are next calculated using a 3<sup>rd</sup> order MLS fit. These results are presented in Figure 9. These results, along with the 1<sup>st</sup> and 2<sup>nd</sup> order MLS fit results, indicate that for randomized calculation points the reducing and smoothing of condition numbers  $C_{max}$  increases with the order of the MLS fit. Furthermore, these results support the statement that the higher order basis functions are more likely to be nearly redundant. In contrast to the 1<sup>st</sup> and 2<sup>nd</sup> MLS order fit results, the minimum number of nearest neighbors  $N$  required to achieve the lowest possible error  $E_{rms}$  is less for randomized calculation points.

The GJE solver results are indistinguishable from those presented for the SVD solver in Figure 9(a) and Figure 9(b) in the region where the condition number is bounded to values  $C_{max} \ll 10^{12}$ , similar to the results for lower order MLS fits. However, in contrast to the previous results, the condition number does not abruptly become unbounded for uniformly spaced calculation points (rather, it quickly grows to approximately  $10^{17}$ ). As a result, the GJE solver does not *abruptly* fail; the error  $E_{rms}$  increases quickly over the range  $40 \leq N \leq 75$  and then fails at  $N < 40$ . Using randomized points, the condition number  $C_{max}$  quickly increases and the GJE solver generates large  $E_{rms}$  errors (for  $14 \leq N \leq 20$ ) and eventually fails as  $C_{max}$  increases without bound. In contrast, the SVD solver again provides solutions with increased  $E_{rms}$  errors for poorly conditioned and ill-conditioned matrices. The difference in behavior between results generated using the two solvers is shown graphically by the selected GJE solver results included in Figures 9(a) and 9(b).

As with lower order MLS fits, the SVD solver provides accurate solutions for a range of poorly conditioned and ill-conditioned matrices for the 3<sup>rd</sup> order fit. If the criterion in setting  $N$  for the 3<sup>rd</sup> order fit is simply that the error  $E_{rms}$  be below  $10^{-2}$  for uniformly spaced or randomly spaced calculation points, then  $N = 40$  is satisfactory. Let the acceptable minimum number of nearest neighbors for a 3<sup>rd</sup> order MLS fit be  $N_{min} = 40$ . If the criterion is again added to avoid solving an ill-conditioned matrix, a value of  $N = 76$  might be chosen. Let the acceptable maximum number of nearest neighbors for a 3<sup>rd</sup> order MLS fit be  $N_{max} = 76$ . As with the 2<sup>nd</sup> order MLS fit, selecting the actual value to be used will be deferred until after review of test results presented in succeeding subsections.



(a)  
uniformly  
spaced  
points



(b)  
randomized  
points  
 $\epsilon = 0.50$

Figure 9. Results for the 3<sup>rd</sup> order MLS fit. The condition number  $C_{max}$  (---□---) and the error  $E_{rms}$  for the gradient (SVD:—○—; GJE:---●---), the Laplacian (SVD:—△—; GJE:---▲---) and the 3<sup>rd</sup> order gradient (SVD:—▽—; GJE:---▼---).

In Figure 9(a), the error in the Laplacian using very few nearest neighbors is seen to be markedly better than that of the 1<sup>st</sup> and 3<sup>rd</sup> order gradients for the case of uniformly spaced calculation points. This experience of an approximation yielding an exceptionally higher order of approximation of a given function occurs in computational approaches employing grids with uniformly spaced calculation points. Such a gift is explained by the “canceling out” of terms of higher order than those calculated in the approximation, producing an effectively higher order approximation. Further details of such occurrences are not within the scope of this MLS investigation and this result will simply be accepted with equanimity.

The value of  $N$  fixes the volume of a sphere, the “MLS window,” circumscribing the  $N$  nearest neighbors. An effective radius  $r$  of this MLS window can be quickly determined from the relation  $N(\Delta x)^3 = \frac{4}{3}\pi r^3$ , where the average calculation point spacing is given by  $\Delta x = \frac{1}{N_s - 1}$ . Hence,

$$r = 0.620 \frac{N^{1/3}}{N_s - 1}. \quad (19)$$

Therefore, the size of MLS window changes by nearly a factor of 4 over the range tested for the 1<sup>st</sup> and 2<sup>nd</sup> order MLS fits and by approximately a factor of 2 for the 3<sup>rd</sup> order fits. The insensitivity of  $E_{rms}$  to the MLS window size is observed in the presented results; with increasing  $N \geq N_{min}$ ,  $E_{rms}$  errors either remain fairly constant (for the 1<sup>st</sup> and 3<sup>rd</sup> order MLS fits) or increase only slightly (for the 2<sup>nd</sup> order fit). As the MLS fits of this section are applied to the test function with a constant value  $\delta = 0.035$ , it follows that  $E_{rms}$  error is also not very sensitive to the ratio of the length scales  $\frac{r}{\delta}$ , where  $\delta$  is used here as a convenient length scale representing the severity of the function curvature.

### Subsection Summary

For the MLS fits tested with either uniformly spaced calculation points or randomized points: (i) there is a minimum number of nearest neighbors  $N_{min}$  such that  $N_{min} \leq N$  maintains error  $E_{rms}$  below  $10^{-2}$  and a maximum number of nearest neighbors  $N_{max}$  such that  $N_{max} \leq N$  in addition avoids an ill-conditioned MLS coefficient matrix even for uniformly spaced calculation points, (ii) the number of nearest neighbors  $N$  in the range  $N_{min} \leq N \leq N_{max}$  increases with increasing order of MLS fit, (iii) using SVD with  $N < N_{min}$  the error  $E_{rms}$  increases either abruptly or quickly by about 2 orders of magnitude greater than the error with  $N_{min} < N$ , and (iv)  $E_{rms}$  errors are nearly insensitive to the size of the MLS window as represented by the ratio of length scales  $\frac{r}{\delta}$ . The condition numbers

$C_{max}$  (v) increase abruptly at  $N = N_{max}$  for uniformly spaced points and increase gradually, “are smoothed”, for randomized points; for orders of MLS fit greater than unity, this occurs at  $N$  considerably less than  $N_{max}$ ; (vi) this smoothing is attributed to the randomized calculation points making possible unique contributions of the basis function to the MLS coefficient matrix. (vii) The values of error  $E_{rms}$  and condition number  $C_{max}$  are not sensitive to the randomization within the range  $0.05 \leq \varepsilon \leq 0.50$ . (viii) The SVD solver does not fail over the tested range  $1 < N < 100$  of nearest neighbors used in the MLS fits. (ix) The GJE solver produces highly inaccurate results for a poorly conditioned MLS coefficient matrix and fails for an ill-conditioned matrix. (x) The  $E_{rms}$  errors are fairly insensitive to the size of the MLS window. (xi) The MLS approach, a representative of residual minimization approaches, calculates values on randomized calculation points with equal or greater accuracy as measured by  $E_{rms}$  and with greater stability as measured by  $C_{max}$  than on uniformly spaced points.

### 3.2.2 Resolution $N_s$ .

The orders of accuracy of the 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> order MLS fits versus  $N_s$  are not 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> order as might be expected. In this subsection, the total number of points  $\bar{N} = N_s^3$  resolving the three-dimensional functions is varied from  $9261 \leq \bar{N} \leq 132651$ , nearly 1.5 orders of magnitude, by varying the number of points per side  $N_s$  of the  $1 \times 1 \times 1$  test space in the range  $21 \leq N_s \leq 51$ .

Using nearest neighbors  $N = 7$ , error  $E_{rms}$  results for the 1<sup>st</sup> order MLS fit are shown in Figure 10. The order of accuracy is strongly dependent on whether the calculation points are uniformly spaced or randomized. The uniformly spaced points show a 2<sup>nd</sup> order accuracy and the randomized points show only a 1<sup>st</sup> order accuracy. The small effect of resolution  $N_s$  on the condition number  $C_{max}$  is shown in Figure 11. The condition number  $C_{max}$  remains constant as the resolution increases using uniformly spaced calculation points and it *increases* slightly as the resolution increases using randomized points.

For the 2<sup>nd</sup> order MLS fit using  $N = N_{min} = 14$ , the order of accuracy for the gradient is again 2<sup>nd</sup> order, shown in Figure 12. The Laplacian shows slightly better than a 1<sup>st</sup> order fit. These results hold for both uniform and randomized calculation points. The condition numbers  $C_{max}$  for the uniform and randomized cases, shown in Figure 13, are fairly constant. However, the condition numbers  $C_{max}$  for the uniformly spaced points are about 15 orders of magnitude greater than those for randomized points, consistent with Figure 8 in the last subsection. Using  $N = N_{max} = 23$ , Figures 14(a) and 14(b) show a somewhat reduced sensitivity to calculation point position randomization: the gradient calculation continues to exhibit nearly 2<sup>nd</sup> order accuracy independent of randomization and only the Laplacian calculation shows sensitivity to randomization with 1.6 order accuracy for uniformly spaced points but only 0.4 order accuracy for randomized points. To see if this order of accuracy trend of decreasing sensitivity to randomization correlates to the increase in the number of nearest neighbors  $N$ , Figure 14(c) is presented on the following page. The single distinction between Figures 14(b) and 14(c) is the increase in  $N$  from 23 to 32. Indeed, for  $N \geq N_{min}$ , the order of accuracy decrease in sensitivity to calculation point randomization is a function of  $N$ .

The condition number values change greatly when varying the number of nearest neighbors used in the 2<sup>nd</sup> order MLS fit from  $N = N_{min} = 14$  to  $N = N_{max} = 23$  as shown in Figure 15. The condition numbers  $C_{max}$  for the uniformly spaced points drop greater than 15 orders of magnitude to a values now *lower* than that for the randomized calculation points. Furthermore, the randomized case now exhibits some, though insignificant, sensitivity to resolution, again *increasing* slightly with increasing resolution.

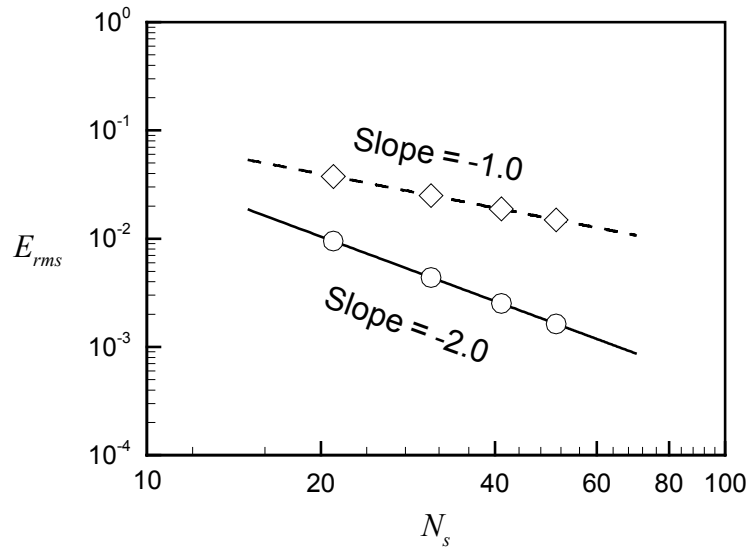


Figure 10. The error  $E_{rms}$  for the gradient with uniformly spaced (—○—) and randomized (---◇---) calculation points,  $\varepsilon = 0.50$ , using a 1<sup>st</sup> order MLS fit. The number of nearest neighbors used is  $N = 7$ .

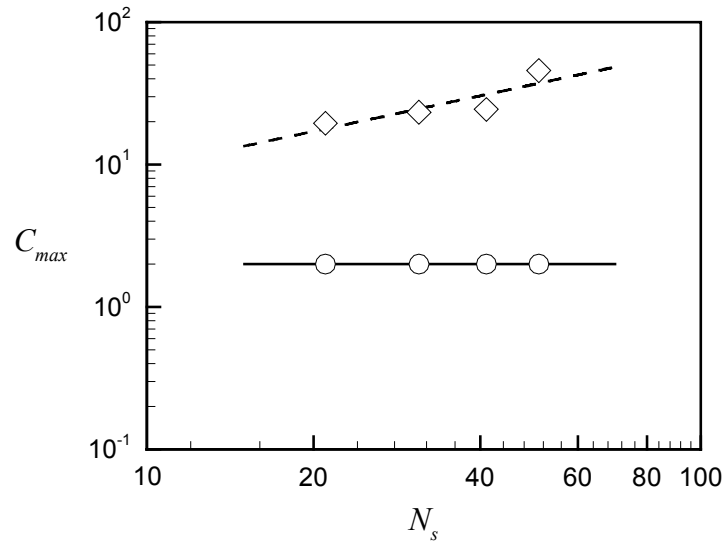


Figure 11. The condition number  $C_{max}$  for the gradient with uniformly spaced (—○—) and randomized (---◇---) calculation points,  $\varepsilon = 0.50$ , using a 1<sup>st</sup> order MLS fit. The number of nearest neighbors used is  $N = 7$ .

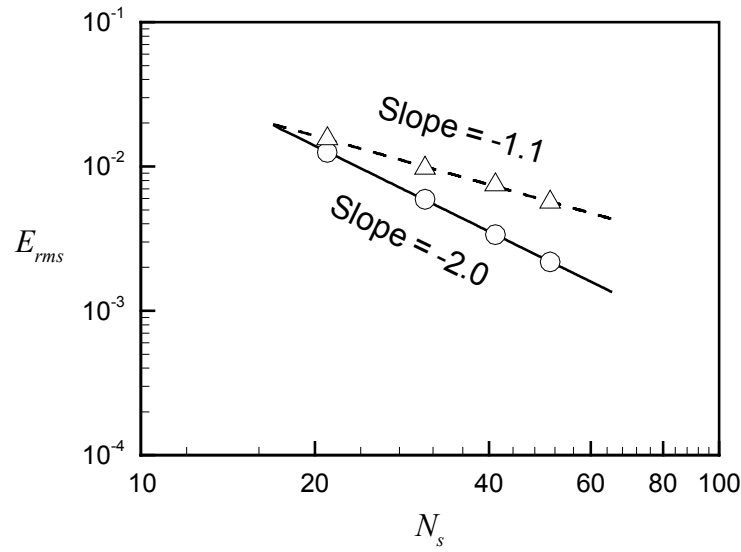


Figure 12. The error  $E_{rms}$  for the gradient (—○—) and Laplacian (---△---) using a 2<sup>nd</sup> order MLS fit. The number of nearest neighbors used is  $N = N_{min} = 14$ . Results apply to calculation points either uniformly spaced or randomized with  $\varepsilon = 0.50$ .

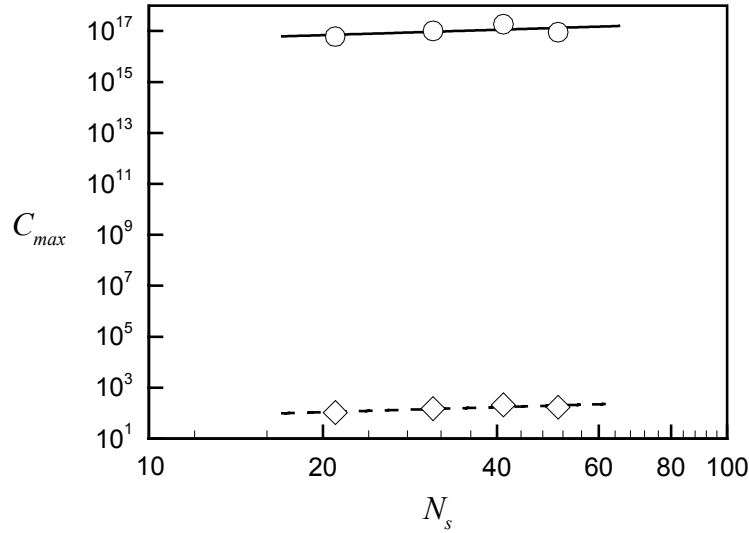
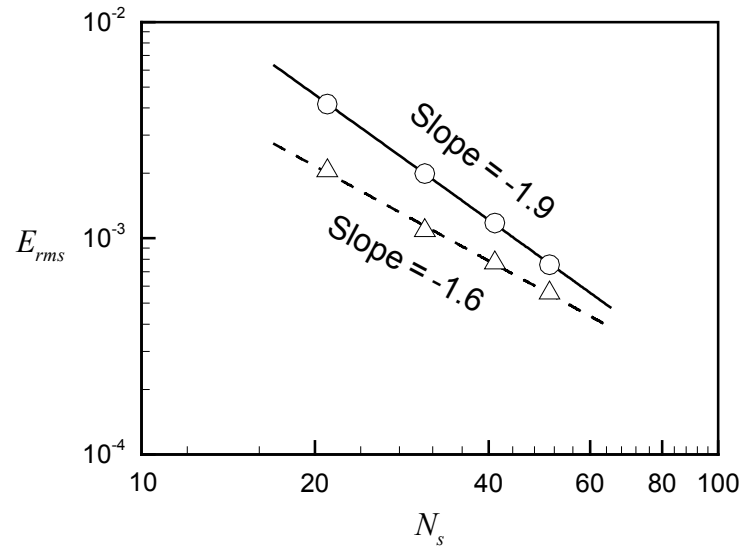
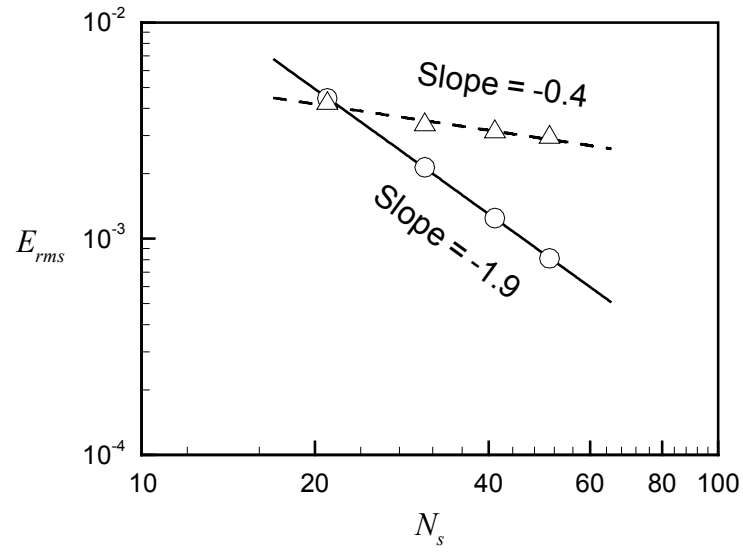


Figure 13. The condition number  $C_{max}$  for uniformly spaced (—○—) and randomized (---◇---) calculation points,  $\varepsilon = 0.50$ , using a 2<sup>nd</sup> order MLS fit. The number of nearest neighbors used is  $N = N_{min} = 14$ .





(a)  
uniformly  
spaced  
points



(b)  
randomized  
points  
 $\epsilon = 0.50$

Figure 14. The error  $E_{rms}$  for the gradient (—○—) and Laplacian (---△---) using a 2<sup>nd</sup> order MLS fit. The number of nearest neighbors used is  $N = N_{max} = 23$ .

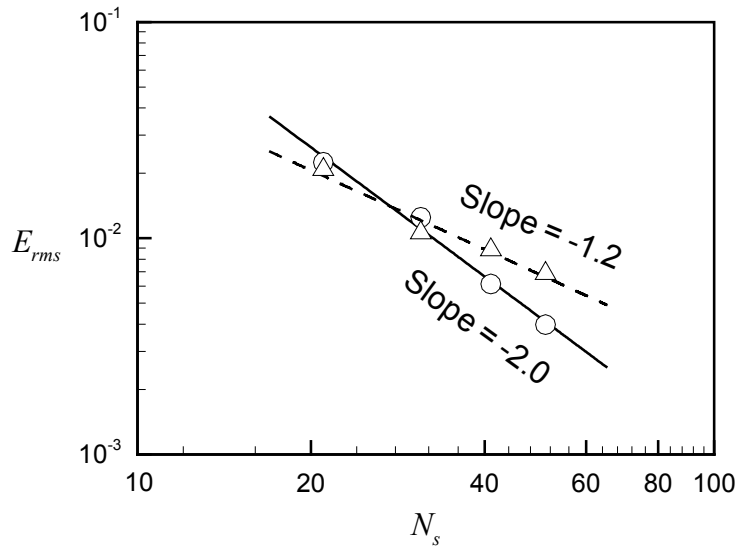


Figure 14(c). The error  $E_{rms}$  for the gradient (—○—) and Laplacian (---△---) using the same parameters as those of Figure 14(b) with one exception: the number of nearest neighbors used is  $N = 32$ .

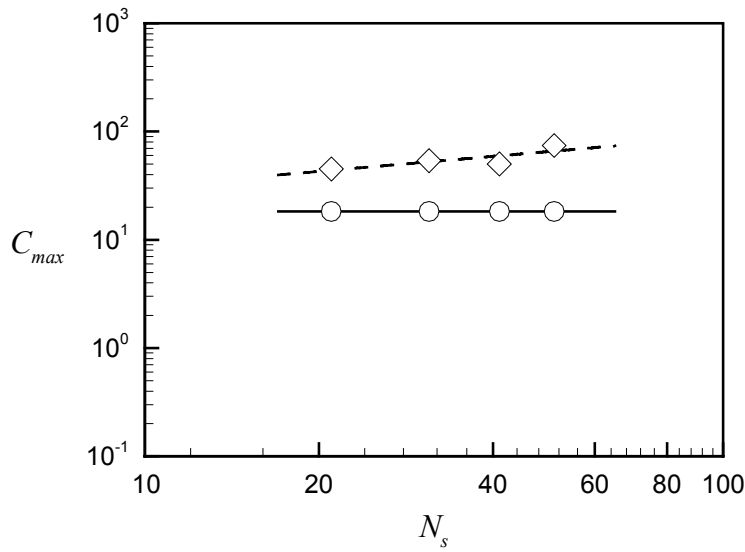
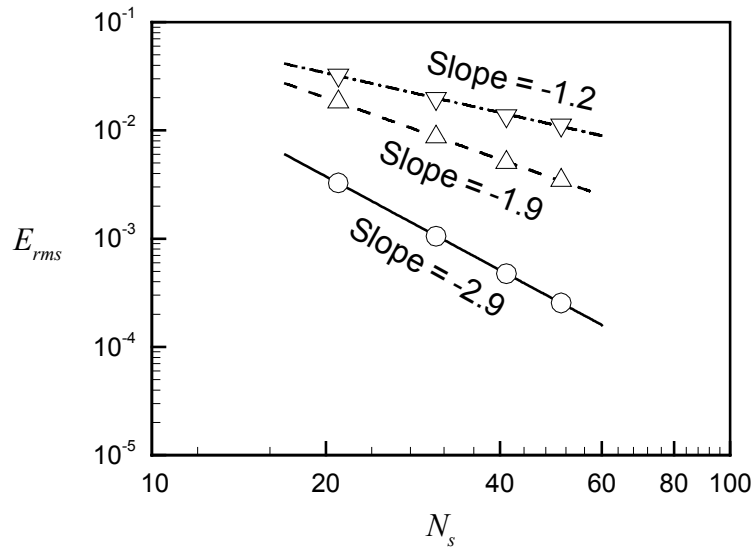


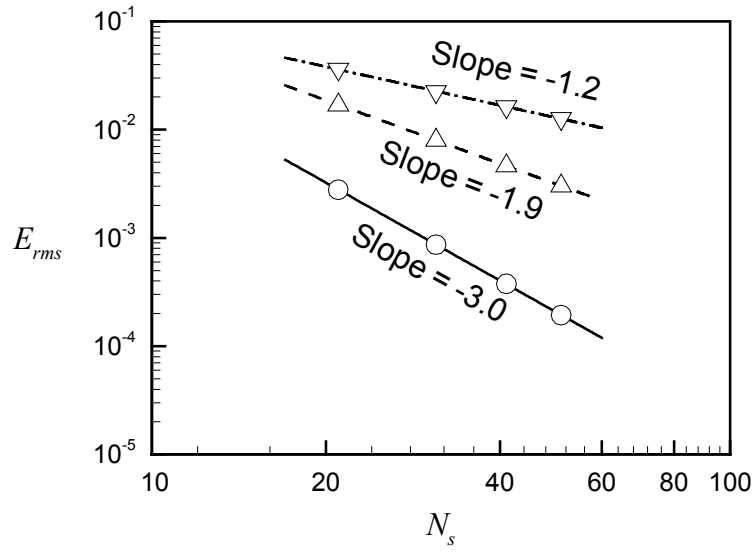
Figure 15. The condition number  $C_{max}$  for uniformly spaced (—○—) and randomized (---◇---) calculation points,  $\varepsilon = 0.50$ , using a 2<sup>nd</sup> order MLS fit. The number of nearest neighbors used is  $N = N_{max} = 23$ .

For the 3<sup>rd</sup> order MLS fit using  $N = N_{min} = 40$ , the order or accuracy has no meaningful sensitivity to whether the calculation points are randomized. Results for the uniformly spaced calculation points and for randomized points are shown in Figure 16. Comparing the 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> order MLS fits, for  $N \geq N_{min}$  the order of accuracy of all derivative calculations generally becomes less sensitive to randomization of calculation point positions as the number of nearest neighbors  $N$  increases. The condition numbers  $C_{max}$  for the uniform and randomized cases, shown in Figure 16, are fairly constant. Similar to the 2<sup>nd</sup> order MLS fit results, the condition numbers for the uniformly spaced calculation points are about 13 orders of magnitude greater than that for the randomized points, shown in Figure 17. This difference in condition numbers  $C_{max}$  is consistent with Figure 9 of the previous subsection. Using  $N = N_{max} = 76$  the error  $E_{rms}$  results change slightly, as shown in Figure 18. For uniformly spaced calculation points, the gradient calculations have a 1.8 order accuracy and the randomized points have a slightly lower 1.5 order accuracy. The Laplacian exhibits nearly 2nd order accuracy independent of calculation point randomization. The 3<sup>rd</sup> order gradient exhibits orders of accuracy of 3.6 and 3.2, respectively, for uniformly spaced and randomized points. The condition number values again change greatly when varying the number of nearest neighbors used in the 3<sup>rd</sup> order MLS fit from  $N = N_{min} = 40$  to  $N = N_{max} = 76$ , as shown in Figure 19. The condition number  $C_{max}$  for the uniform case drops about 14 orders of magnitude to values now *lower* than that for the randomized calculation point case; in this 3<sup>rd</sup> order MLS fit, however,  $C_{max}$  does not exhibit sensitivity to resolution for either uniformly spaced or randomized calculation points.

As discussed in Subsection 3.2.1 (and will be further discussed in Subsection 3.2.4),  $E_{rms}$  error is not very sensitive to the ratio of the length scales  $\frac{r}{\delta}$  including those corresponding to the changes in resolution  $N_s$ . For all MLS fits tested, the order of accuracy is, therefore, attributed to the relative changes in the average spacing  $\Delta x = \frac{1}{N_s - 1}$  given by  $\frac{\Delta x}{\delta}$ , where  $\delta$  is used here as a convenient length scale representing the severity of the function curvature.



(a)  
uniformly  
spaced  
points



(b)  
randomized  
points  
 $\epsilon = 0.50$

Figure 16. The error  $E_{rms}$  for the gradient (—○—), Laplacian (---△---) and 3<sup>rd</sup> order gradient (---▽---) using a 3<sup>rd</sup> order MLS fit. The number of nearest neighbors used is  $N = N_{min} = 40$ .

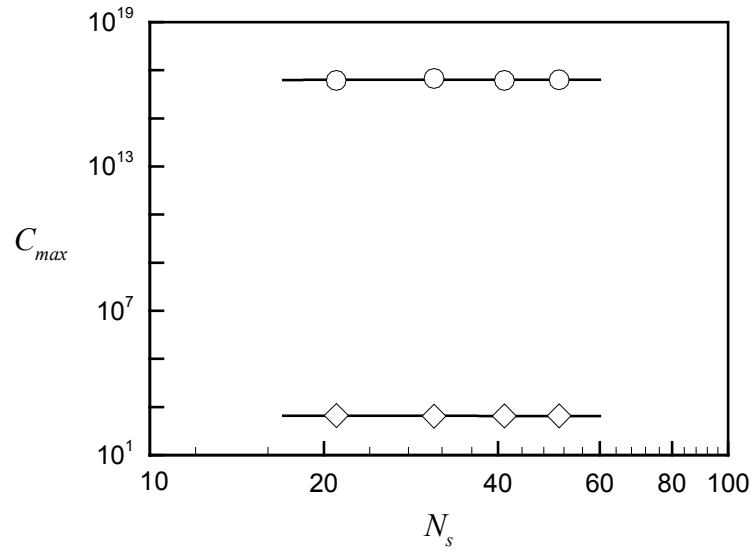


Figure 17. The condition number  $C_{max}$  for uniformly spaced (—○—) and randomized (---◇---) calculation points,  $\varepsilon = 0.50$ , using a 3<sup>rd</sup> order MLS fit. The number of nearest neighbors used is  $N = N_{min} = 40$ .

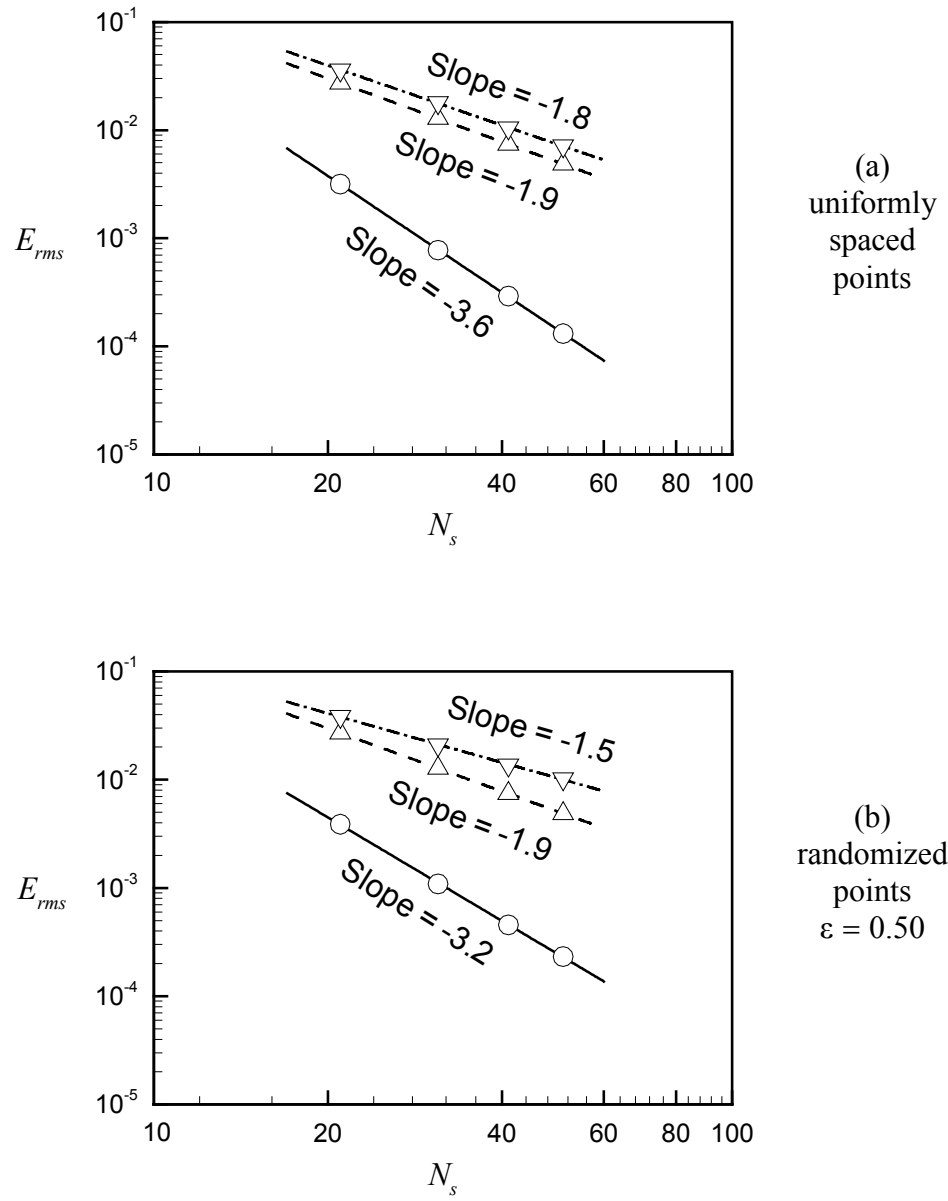


Figure 18. The error  $E_{rms}$  for the gradient (—○—), Laplacian (---△---) and 3<sup>rd</sup> order gradient (---▽---) using a 3<sup>rd</sup> order MLS fit. The number of nearest neighbors used is  $N = N_{max} = 76$ .

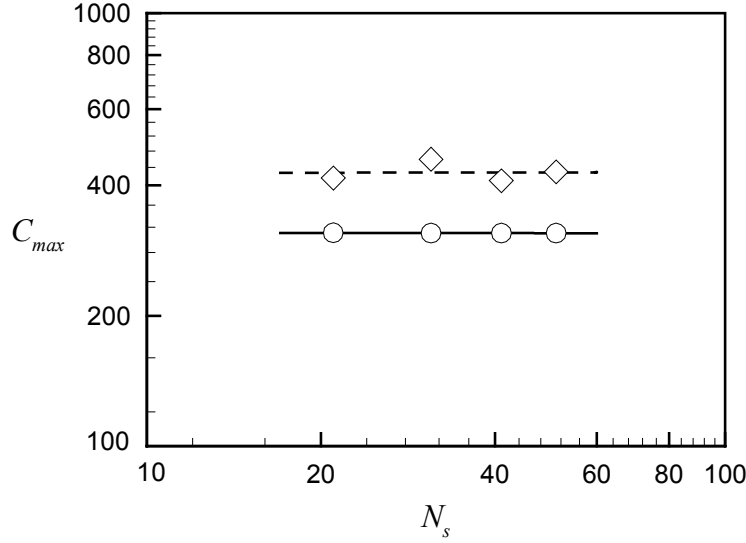


Figure 19. The condition number  $C_{max}$  for uniformly spaced (—○—) and randomized (---◇---) calculation points,  $\varepsilon = 0.50$ , using a 3<sup>rd</sup> order MLS fit. The number of nearest neighbors used is  $N = N_{max} = 76$ .

### Subsection Summary

(i) For an  $h^{\text{th}}$  order MLS fit, the nominal order of accuracy for approximation of a  $j^{\text{th}}$  order gradient is roughly  $(h - j + 1)$ . (ii) The observed orders of accuracy are due to the ratio of changes of average point spacing to a measure of the severity of function curvature  $\frac{\Delta x}{\delta}$ . (iii)

A broad trend is observed in which the order of accuracy of derivative calculation becomes less sensitive to randomization of points as the number of nearest neighbors  $N$  used in the MLS fits increases. (iv) The order of accuracy is either unchanged or slightly improved for an MLS fit using either (a) more nearest neighbors in the range  $N_{min} \leq N \leq N_{max}$  or (b) uniformly spaced calculation points; (v) the condition number  $C_{max}$  exceeds  $10^{16}$  indicating the coefficient matrix is ill-conditioned for the 2<sup>nd</sup> and 3<sup>rd</sup> order MLS fits using uniformly spaced calculation points and  $N_{min}$  nearest neighbors; (vi) the condition numbers  $C_{max}$  are bounded to  $C_{max} \ll 10^{12}$  and are almost constant for an MLS fit using randomized calculation point locations.

### 3.2.3 Scaled isolation distance $D$ .

Inspection of the maximum condition number  $C_{max}$  used throughout this report shows that this maximum consistently occurs at the “corners” of the  $1 \times 1 \times 1$  test volume. That is, the condition number is greatest at the “most isolated” points. These corner points also have a one-sided symmetry. The isolation of and symmetry about a point are explored by investigating the error  $E_{rms}$  and the maximum condition number  $C_{max}$  for various isolated points with and without function symmetry about them. “Externally” isolated points and “internally” isolated points are investigated using the test arrangements depicted, respectively, in Figures 20 and 21. Shown in each of these figures are uniformly spaced “bulk” calculation points and a single isolated point placed “on-center” in a location of symmetry. The bulk calculation points have an average local spacing  $\Delta x$ ; the isolated points do not. For isolated points the “scaled isolation distance” is used, defined as the distance between an isolated calculation point and the nearest neighbor. This distance is measured as a multiple  $D$  times the average bulk calculation point spacing  $\Delta x$ . The isolated points in Figures 20 and 21 are shown at a scaled isolation distance  $D = 5$ . The number and arrangement of bulk calculation points used in isolated external point tests is a  $16 \times 31 \times 31$  block of points and for the isolated internal point tests the number and arrangement of the bulk calculation points is  $31 \times 31 \times 31$  with a  $(2D - 1)^3$  block of points removed. In various indicated tests, all points other than the isolated point are also randomized with maximum perturbation  $\varepsilon$  as described earlier in this report. Isolated external points termed “off-center” are not placed in positions of symmetry; they are offset from the  $y = 0.5$  line by a distance  $D$  along the  $x$ -axis and  $2D$  along the  $z$ -axis. Isolated internal points not placed in positions of symmetry are offset from the point  $(x, y, z) = (0.5, 0.5, 0.5)$  by a distance  $D$  along the  $x$ -axis and  $2D$  along the  $y$ -axis and  $3D$  along the  $z$ -axis.

All tests in this subsection use a 2<sup>nd</sup> order MLS fit with resolution given by  $N_s = 31$ . As with previous tests, the SVD solver is used except where indicated.

Figures 22 and 23 show the condition number  $C_{max}$  and error  $E_{rms}$  for the gradient and Laplacian calculations with the number of nearest neighbors  $N = N_{max} = 23$  used in the MLS fit. The isolated points in these figures are external, either on-center or off-center, and the calculation points are uniformly spaced. The scales of these figures are the same as those of related figures presented later to provide for direct comparison. Results show the condition number sharply increases as the scaled isolation distance  $D$  increases beyond the value of unity. The error  $E_{rms}$  remains fairly constant at about  $10^{-2}$ . Using the GJE solver (results not shown) rather than the SVD solver for the on-center calculations shown in Figure 22, yields  $E_{rms}$  errors of the same value at  $D = 1$ . As the condition number  $C_{max}$  sharply increases, the  $E_{rms}$  errors grow for the GJE calculated values (to approximately  $4 \times 10^{-2}$  at  $D = 3$ ) and then fails (at  $D \geq 4$ ).



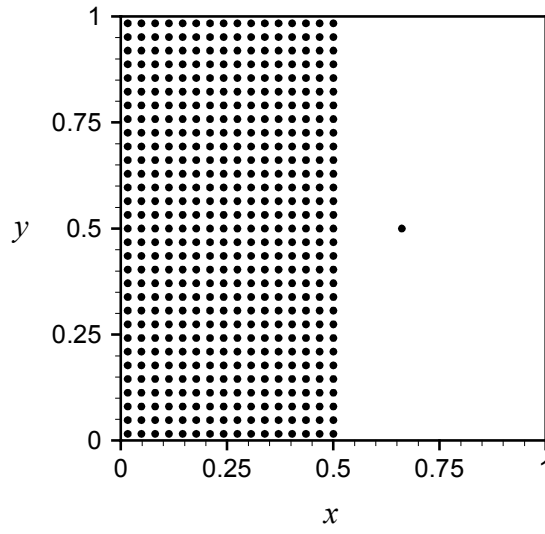


Figure 20. Depiction of the external centered isolated point showing uniformly spaced bulk calculation points and scaled isolation distance  $D = 5$ .

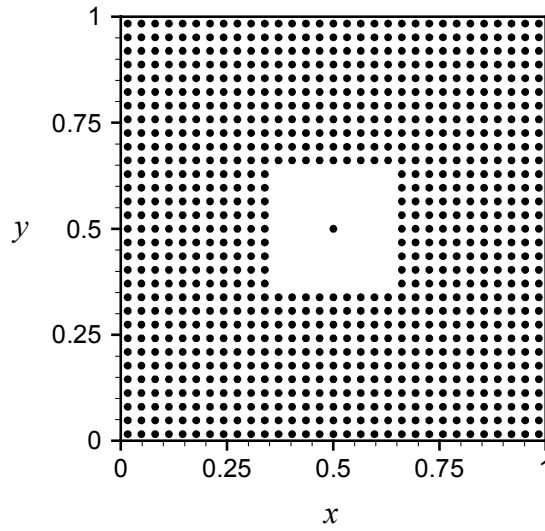


Figure 21. Depiction of the internal centered isolated point showing uniformly spaced bulk calculation points and scaled isolation distance  $D = 5$ .

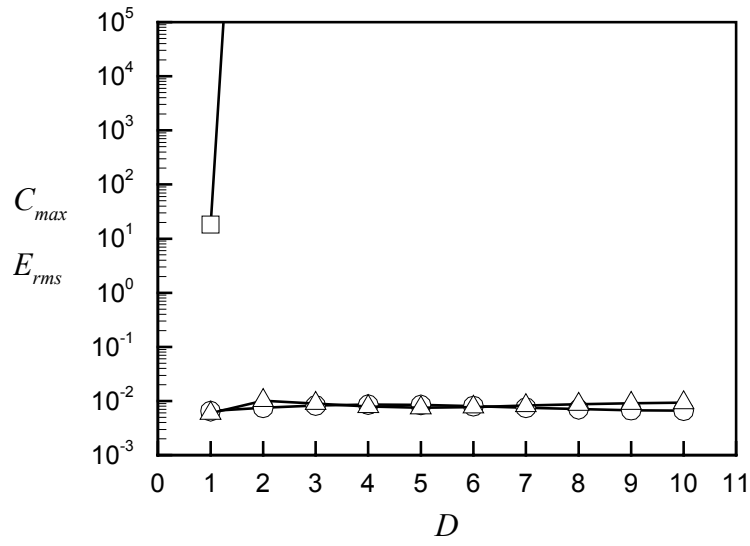


Figure 22. The condition number  $C_{max}$  (—□—) and error  $E_{rms}$  for the gradient (—○—) and Laplacian (—△—) with  $N = N_{max} = 23$ . Bulk calculation points are uniformly spaced. The isolated point is external, on-center.

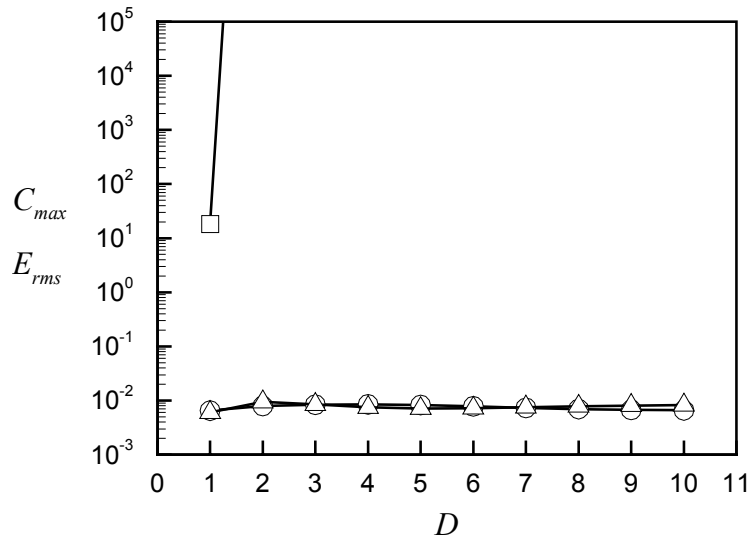


Figure 23. The condition number  $C_{max}$  (—□—) and error  $E_{rms}$  for the gradient (—○—) and Laplacian (—△—) with  $N = N_{max} = 23$ . Bulk calculation points are uniformly spaced. The isolated point is external, off-center.

The results for the on-center or off-center isolated points are in close agreement. This agreement is observed in all external and internal isolated point tests considered supporting a conclusion that function symmetry does not contribute to  $C_{max}$  in these or previous tests. Henceforward, only the symmetrical on-center test results will be shown and these will be discussed simply as results of isolated point tests, it being understood that the presented results are very similar to those of the off-center tests.

Results for bulk calculation points greatly randomized ( $\varepsilon = 0.50$ ) or minimally randomized ( $\varepsilon = 0.05$ ) are shown respectively in Figures 24 and 25. In these figures, error  $E_{rms}$  values are similar in magnitude (though the errors in the gradient and Laplacian calculations switch places) to those for the uniformly spaced bulk points;  $E_{rms}$  is not sensitive to the degree of randomization. Consistent with previous tests in this report, the randomization of the bulk calculation points greatly reduces the condition number values with the result that  $C_{max}$  values become bounded to  $C_{max} \ll 10^{12}$  over the tested range of  $D$ .

The effects of varying the scaled isolation distance  $D$  when using a reduced number of nearest neighbors  $N = N_{min} = 14$  are shown in Figure 26. There is a small increase in the error  $E_{rms}$  of the gradient and the Laplacian calculation for uniform or randomized bulk calculation points. The condition numbers of the MLS coefficient matrix become ill-conditioned for uniformly spaced calculation points beginning at scaled isolation distance  $D = 1$  using  $N = N_{min} = 14$  nearest neighbors rather than  $D > 1$  using  $N = N_{max} = 23$ ; this is a distinction of trivial practical importance. Using randomized calculation points, no meaningful changes in the condition numbers  $C_{max}$  are evident for  $N = N_{min} = 14$  versus  $N = N_{max} = 23$ .

Using uniformly spaced points and  $N = N_{max} = 23$ , results for the internally isolated point are shown in Figure 27. The condition number is constant with  $C_{max} = 18$ . The error fluctuates somewhat with changing  $D$  but never departs appreciably from  $E_{rms} = 10^{-2}$ . For randomized bulk calculation points, the condition number and error results are fairly insensitive to randomization of the points for internally isolated points.

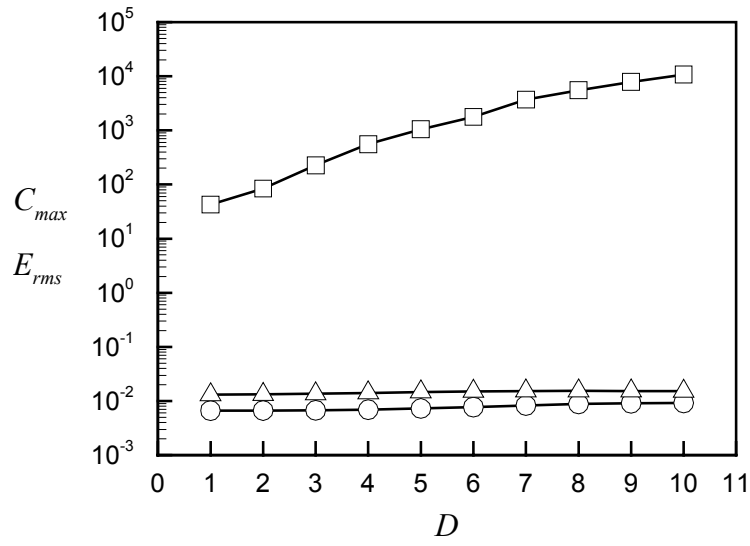


Figure 24. The condition number  $C_{max}$  (—□—) and error  $E_{rms}$  for the gradient (—○—) and Laplacian (—△—) with  $N = N_{max} = 23$ . Bulk calculation points are randomized,  $\varepsilon = 0.50$ . The isolated point is external.

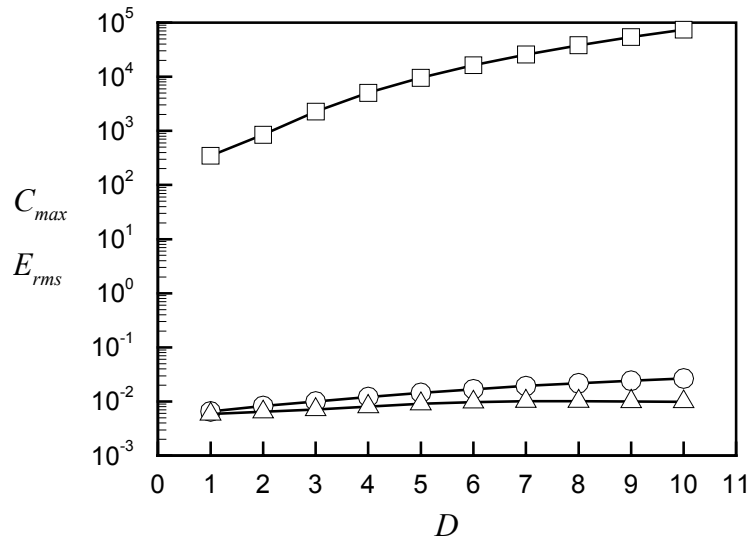
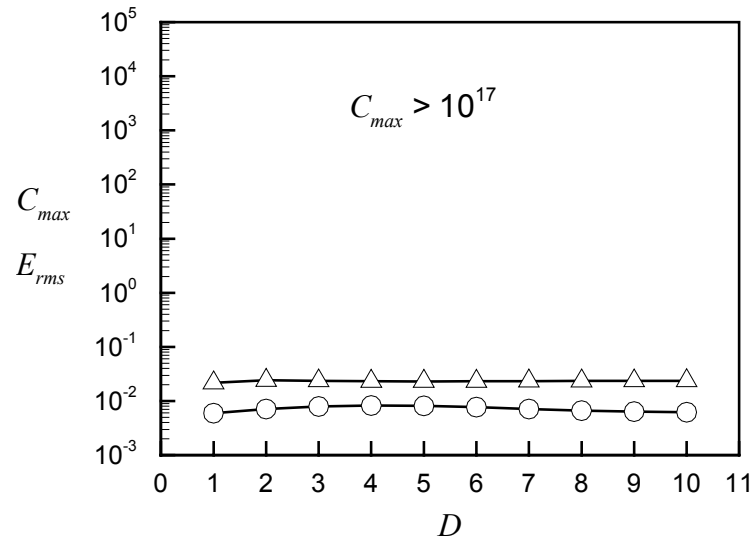
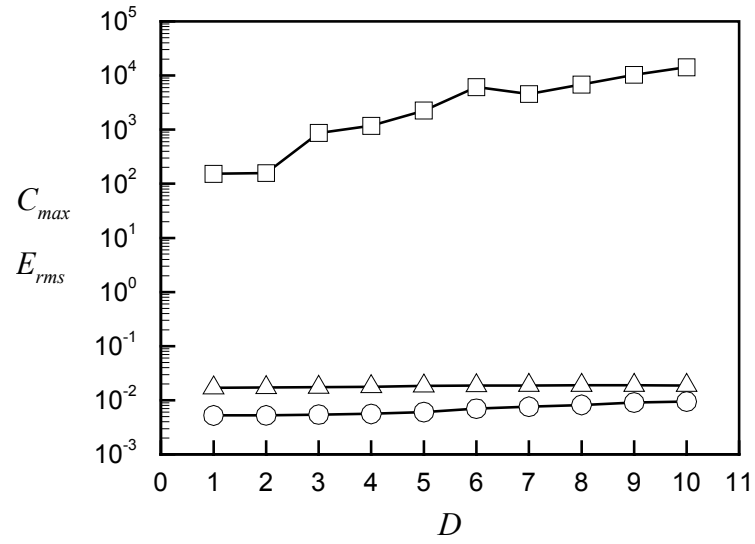


Figure 25. The condition number  $C_{max}$  (—□—) and error  $E_{rms}$  for the gradient (—○—) and Laplacian (—△—) with  $N = N_{max} = 23$ . Bulk calculation points are randomized,  $\varepsilon = 0.05$ . The isolated point is external.

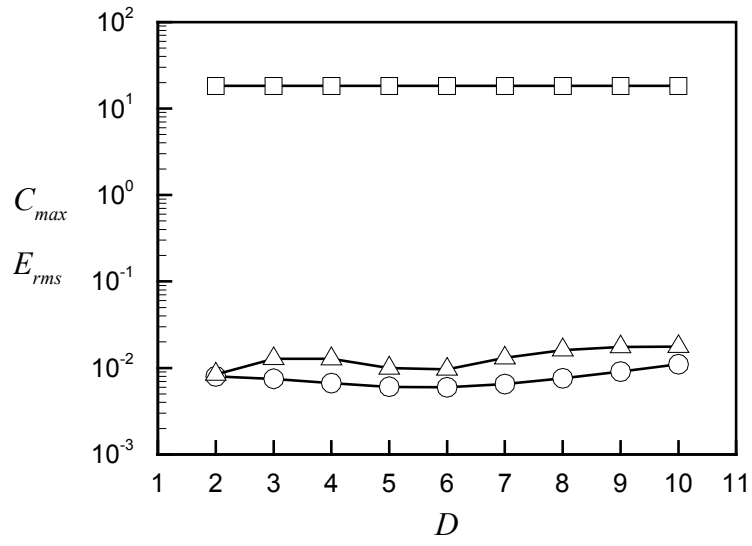


(a)  
uniformly  
spaced bulk  
points

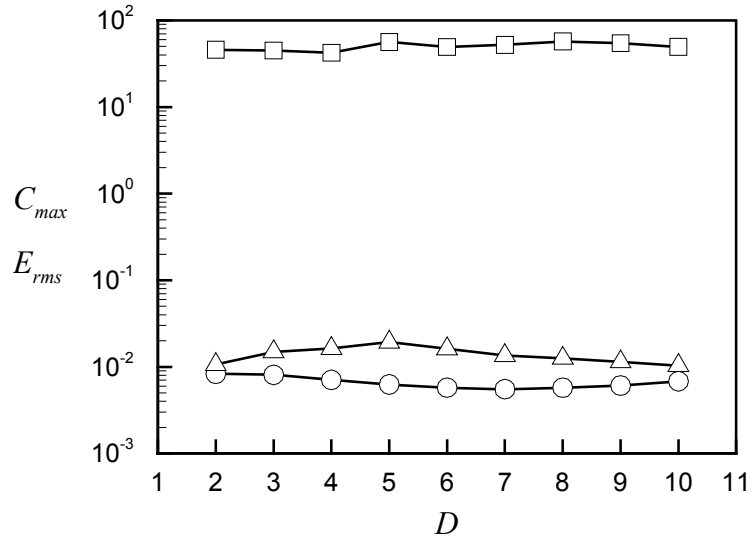


(b)  
randomized  
bulk points  
 $\epsilon = 0.50$

Figure 26. The condition number  $C_{max}$  (—□—) and error  $E_{rms}$  for the gradient (—○—) and Laplacian (—△—) with  $N = N_{min} = 14$ . The isolated point is external.



(a)  
uniformly  
spaced bulk  
points



(b)  
randomized  
bulk points  
 $\epsilon = 0.50$

Figure 27. The condition number  $C_{max}$  (—□—) and error  $E_{rms}$  for the gradient (—○—) and Laplacian (—△—) with  $N = N_{max} = 23$ . The isolated point is internal.

### Subsection Summary

Over a wide range of scaled isolation distances  $1 \leq D \leq 10$  : (i) for uniformly spaced bulk calculation points and any externally isolated point, the MLS coefficient matrix should be considered ill-conditioned; (ii) there is no practical difference in the magnitude or trend of the condition numbers  $C_{max}$  and only a small increase in errors  $E_{rms}$  when isolated points are present, with one exception—for externally isolated points, the condition numbers  $C_{max}$  are highly sensitive to whether bulk calculation point locations are uniformly spaced or randomized, while for internally isolated points the condition numbers are bounded to  $C_{max} \ll 10^{12}$  and nearly independent of randomization; (iii) for isolated points, the error  $E_{rms}$  is not sensitive to whether bulk calculation points are spaced uniformly or randomly; (iv) for isolated points, the error  $E_{rms}$  and the condition number  $C_{max}$  are not sensitive to the degree of randomization  $\varepsilon$ , nor to the number of nearest neighbors within the range  $N_{min} \leq N \leq N_{max}$ , nor to whether there is function symmetry about the point.

### 3.2.4 Gaussian Radius $\delta$ .

Decreasing the Gaussian diameter  $\delta$  given by equation (17) narrows the test function  $f$  as shown in Figure 28 and narrows higher orders of the gradients of  $f$  exponentially. The ratio of the “wave amplitude” to “wave width”  $\Omega$  measures the severity of the curves in these functions and is summarized in Table 3.

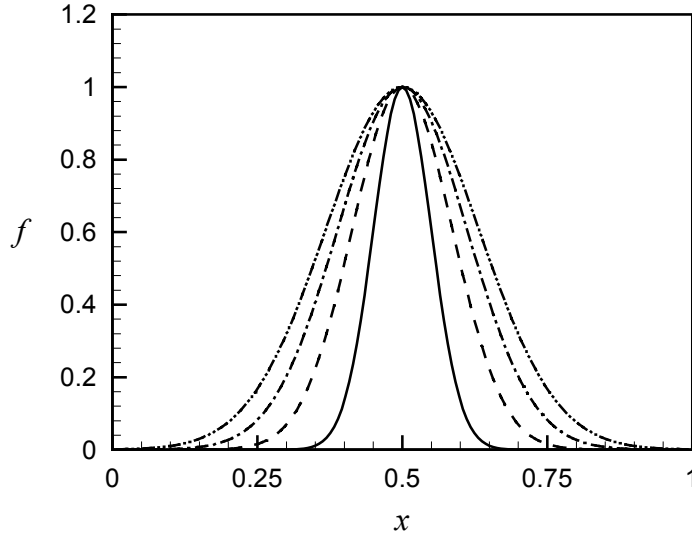


Figure 28. Test function values along the  $x$ -axis for  $\delta^2 = 0.005$  (—),  $\delta^2 = 0.015$  (-----),  $\delta^2 = 0.025$  (.....) and  $\delta^2 = 0.035$  (-·-·-·-).

	$f$	$\nabla f$	$\nabla^2 f$
$\delta^2 = .005$	4	24	4800
$\delta^2 = .015$	2	15	1200
$\delta^2 = .025$	1.5	12	480
$\delta^2 = .035$	1	10	325

Table 3. Severity of function curves measured by the ratio  $\Omega$  of wave amplitude to width.



A 2<sup>nd</sup> order MLS fit using the SVD solver with resolution given by  $N_s = 31$  and with number of nearest neighbors  $N = N_{max} = 23$  is used to study the effects of varying  $\delta$  in the tests that follow. In all tests, results are presented for uniformly spaced calculation points. The gradient results are first reviewed. Beginning with  $\delta^2 = 0.035$ , Figure 29 shows very good agreement between the analytical and calculated solutions.

The condition numbers are constant with  $C_{max} = 18$  for the series of tests shown in Figures 29 through 32. This value is the same as that of the tests presented previously: the condition numbers are not sensitive to the value of  $\delta$  or equivalently the value of  $\Omega$  over the range tested.

Considering  $\delta^2 = 0.025$ , a slight smoothing of the function peaks, the region of greatest curvature, is evident in Figure 30. This behavior becomes more evident in Figure 31 with  $\delta^2 = 0.015$  and Figure 32 with  $\delta^2 = 0.005$  in large part because as the radius decreases the peak amplitude of the derivative function increases; the error  $E_{rms}$  is an rms error scaled to this amplitude and is not necessarily increasing and cannot be well estimated by eye. The results of  $E_{rms}$  error versus the radius  $\delta$  are presented later. In Figure 32, close inspection shows that in addition to the peaks being smoothed and reduced in magnitude, the widths of the curves are being smoothed and broadened.

The Laplacian calculation results are presented for the same values of  $\delta$  in Figures 33 through 36. The trend of peak reduction is plainly evident and that of width broadening is discernable. As for the 1<sup>st</sup> order gradient results, the scaled error  $E_{rms}$  cannot be well estimated by eye in these figures and results of  $E_{rms}$  error versus the radius  $\delta$  are presented later.

For comparison of these results, a maximum error  $E_{max}$  is defined as

$$E_{max} = \text{MAX} \left\{ \frac{\text{ABS} [g'(\bar{\mathbf{x}}_i) - g(\bar{\mathbf{x}}_i)]}{g(\bar{\mathbf{x}}_i)} \right\}. \quad (20)$$

where  $g$  and  $g'$  represent the analytic and calculated values, respectively, of  $\nabla f$  or  $\nabla^2 f$ . The decrease in errors  $E_{rms}$  and  $E_{max}$  with the increase in the Gaussian test function radius  $\delta$  is summarized in Figure 37 for uniformly spaced calculation points in graph (a) and for randomized calculation points in graph (b). There is little sensitivity to  $\delta$ ; the change in  $E_{rms}$  error with decreasing  $\delta$  is nearly zero in all cases. For randomized calculation point locations, the error in the 2<sup>nd</sup> order MLS fit for the Laplacian calculation actually decreases marginally with decreasing  $\delta$ . For the tests in this section, the representative radius  $r$  of the

MLS window size, cf. equation (19), is a constant value ( $r = 0.058773$ ) set by the values  $N = 23$  and  $N_s = 31$ ; hence, these results agree with those of Subsection 3.2.1, further supporting the conclusion that there is little sensitivity of  $E_{rms}$  error to the relative size of the MLS window as represented by the ratio of the length scales  $\frac{r}{\delta}$ , where  $r$  is the effective radius of MLS window and  $\delta$  is used as a convenient length scale representing the severity of the function curvature. The error  $E_{max}$ , however, shows sensitivity to  $\frac{r}{\delta}$  and, with the chosen value of  $N = 23$ , to whether or not the calculation points are randomized. All calculated  $E_{max}$  errors were found to be located at locations of highest curvature, as expected. Hence, the errors associated with a function shape are localized to regions of high curvature and result from the approximation reducing the severity of the curvature.

The changes in error that can be most evidently observed in  $E_{max}$  are a result of the changes in the dimensionless ratio  $\frac{\Delta x}{\delta}$ , the ratio of average point spacing to a convenient measure of the severity of function curvature. These results are consistent with the resolution test results of Subsection 3.2.2 in which observed orders of accuracy are attributed to this same ratio of length scales.

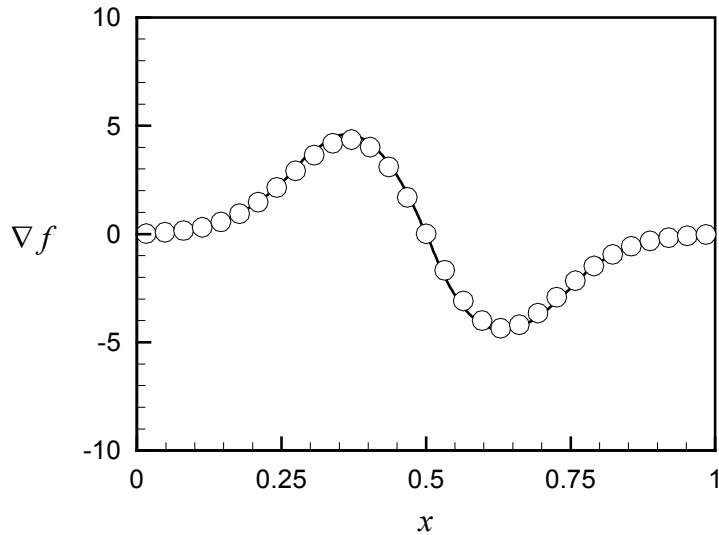


Figure 29. The analytical solution (—) on the  $x$ -axis and a thin center slice of the gradient values calculated ( $\circ$ ) using a 2<sup>nd</sup> order MLS fit on the test function  $f$  with  $\delta^2 = 0.035$ .

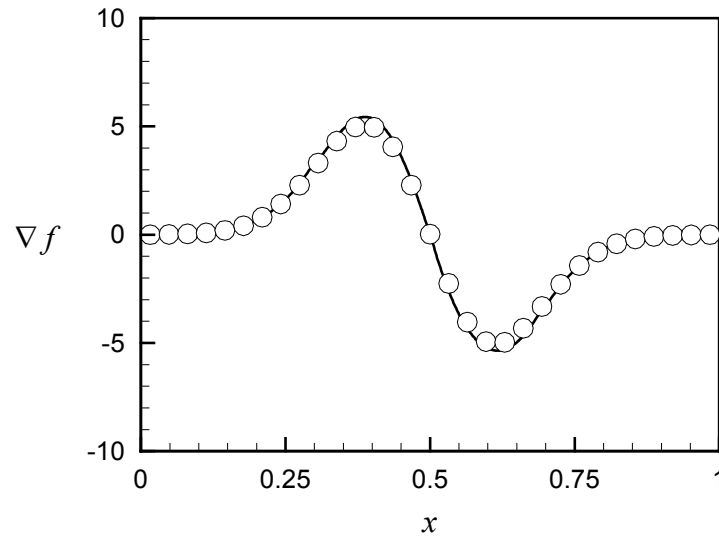


Figure 30. The analytical solution (—) on the  $x$ -axis and a thin center slice of the gradient values calculated ( $\circ$ ) using a 2<sup>nd</sup> order MLS fit on the test function  $f$  with  $\delta^2 = 0.025$ .

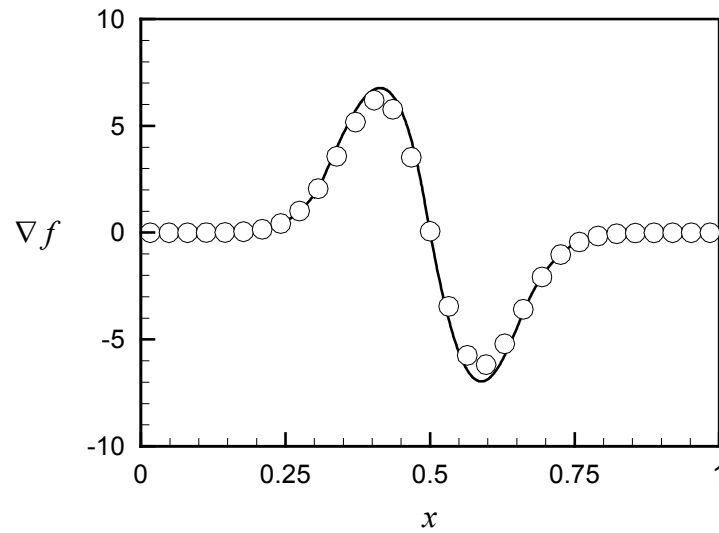


Figure 31. The analytical solution (—) on the  $x$ -axis and a thin center slice of the gradient values calculated ( $\circ$ ) using a 2<sup>nd</sup> order MLS fit on the test function  $f$  with  $\delta^2 = 0.015$ .

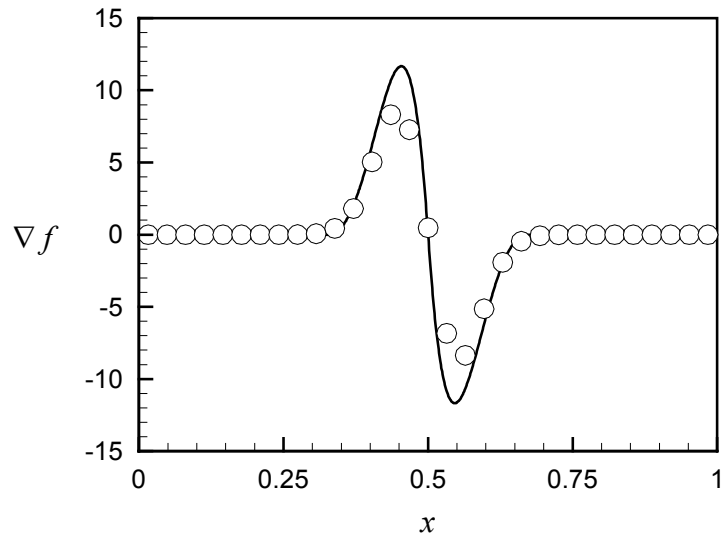


Figure 32. The analytical solution (—) on the  $x$ -axis and a thin center slice of the gradient values calculated ( $\circ$ ) using a 2<sup>nd</sup> order MLS fit on the test function  $f$  with  $\delta^2 = 0.005$ .

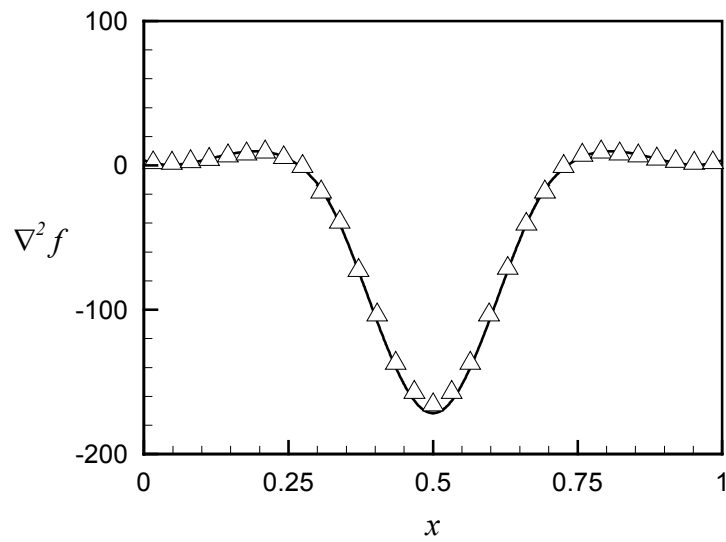


Figure 33. The analytical solution (—) on the  $x$ -axis and a thin center slice of the Laplacian values calculated ( $\triangle$ ) using a 2<sup>nd</sup> order MLS fit on the test function  $f$  with  $\delta^2 = 0.035$ .

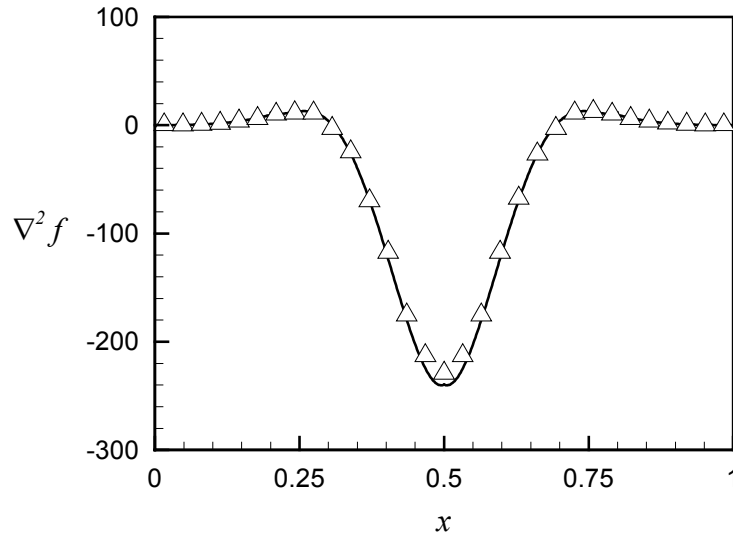


Figure 34. The analytical solution (—) on the  $x$ -axis and a thin center slice of the Laplacian values calculated ( $\triangle$ ) using a 2<sup>nd</sup> order MLS fit on the test function  $f$  with  $\delta^2 = 0.025$ .

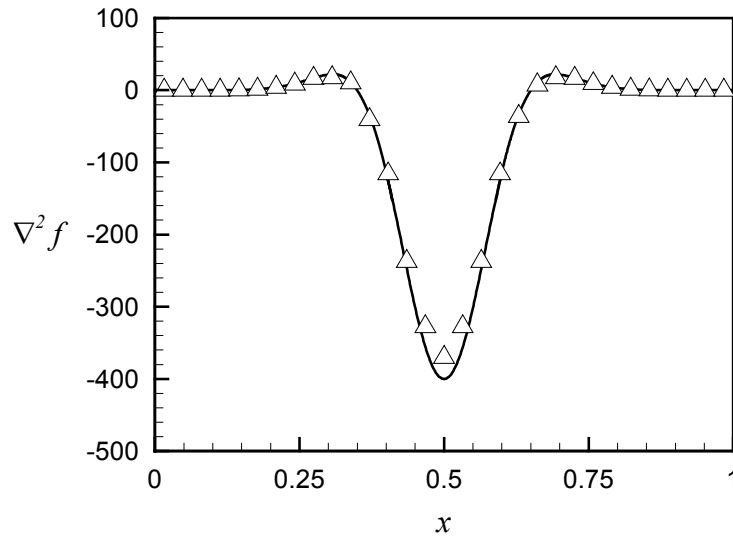


Figure 35. The analytical solution (—) on the  $x$ -axis and a thin center slice of the Laplacian values calculated ( $\triangle$ ) using a 2<sup>nd</sup> order MLS fit on the test function  $f$  with  $\delta^2 = 0.015$ .

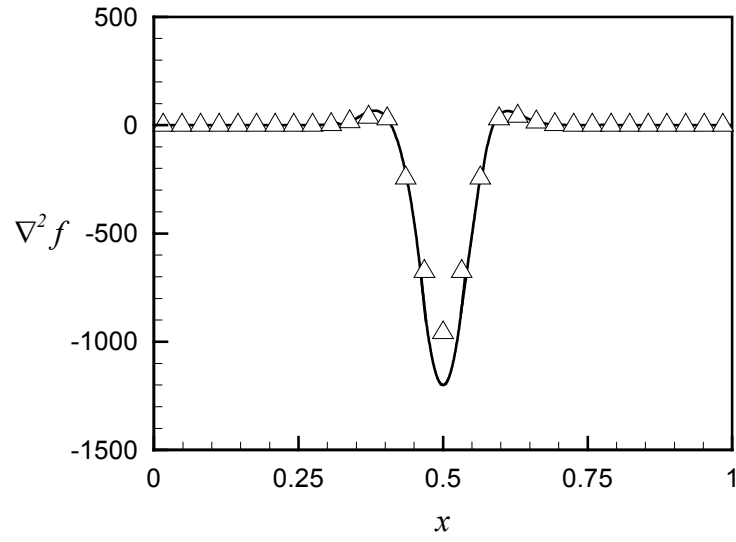
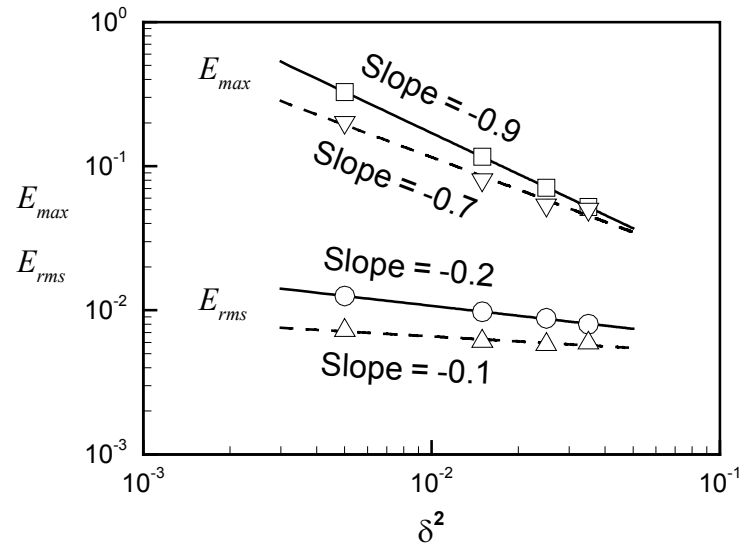
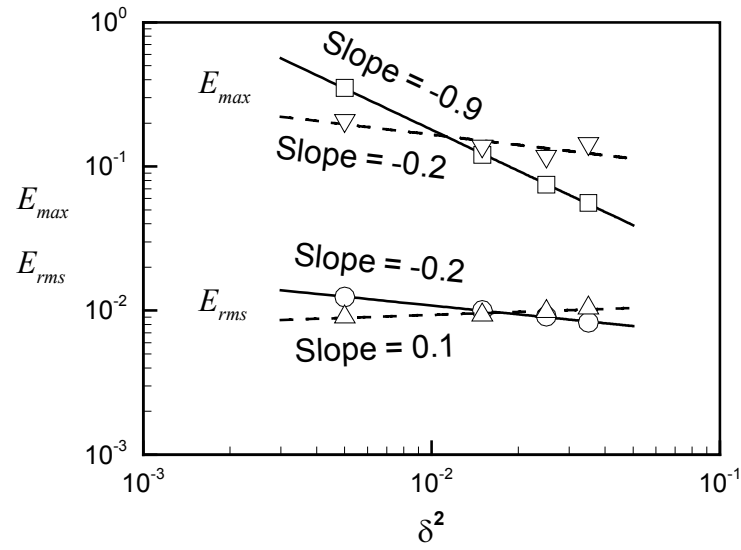


Figure 36. The analytical solution (—) on the  $x$ -axis and a thin center slice of the Laplacian values calculated ( $\triangle$ ) using a 2<sup>nd</sup> order MLS fit on the test function  $f$  with  $\delta^2 = 0.005$ .



(a)  
uniformly  
spaced  
points



(b)  
randomized  
points  
 $\epsilon = 0.50$

Figure 37. Errors in the gradient calculation ( $E_{max}$ :—□—;  $E_{rms}$ :—○—) and in the Laplacian calculation ( $E_{max}$ :---▽---;  $E_{rms}$ :---△---) using a 2<sup>nd</sup> order MLS fit.

### Subsection Summary

(i) Error  $E_{max}$  increases with increasing severity of function curvature and (ii)  $E_{max}$  tends to be localized to regions of high function curvature with the approximation reducing the severity of the curvature; (iii) because the  $E_{max}$  error may be very localized,  $E_{rms}$  may be only weakly dependent on function curvature.



### 3.2.5 Execution Time

The algorithm to perform MLS calculations can conveniently be divided into a nearest neighbors list creation module, independent of the solver used, and a coefficient matrix solver module. The execution times of these algorithm modules are presented in this subsection. All timing tests are performed on a dedicated SUN ultra-2 workstation. A 2<sup>nd</sup> order MLS fit and randomized calculation points with  $\varepsilon = 0.50$  are used in all cases tested in this subsection. All results are presented in seconds.

In the nearest neighbors list creation module for this report, calculation points are first sorted into lists in each Cartesian direction using a binary search procedure. The number of numerical operations per calculation point in this sort is proportional to  $\log \bar{N}$  [16] such that the total number of operations is proportional to  $\bar{N} \log \bar{N}$  (equivalent to order 1.16 for  $10^2 \leq \bar{N} \leq 10^6$ ). Timing results presented in Figure 38 are consistent with this estimate of numerical operations. The spatially sorted lists are next used in creating sorted nearest neighbors lists about each calculation point. The total run time required to create these nearest neighbors lists using  $N = 14$  is shown in Figure 39. The total run time is proportional to  $\bar{N}(\log \bar{N})^3$ ; hence, the number of numerical operations per calculation point is proportional to  $(\log \bar{N})^3$ .

Within the solver module, the MLS coefficient matrix must be solved once for each calculation point. This leads directly to the results shown for the SVD and the GJE solvers versus the total number of points  $\bar{N}$  in Figure 40; the total run time is proportional to  $\bar{N}$  but the SVD solver requires approximately 4 times the total run time of the GJE solver.

Now, to come full circle and fully satisfy one of the opening sentences for the application tests in this report, “The first (and last) consideration in this application is choosing the number of nearest neighbors  $N$ ,” the execution time versus  $N$  is examined. A resolution given by  $N_s = 21$  ( $\bar{N} = 9261$ ) and randomized calculation points with  $\varepsilon = 0.50$  are used in the following tests. The calculation points sorting time is independent of the number of nearest neighbors  $N$  and results of Figure 38 still apply. The total run time to create the nearest neighbors list is proportional to  $\frac{N}{2}$  as shown in Figure 41. This is a small price to pay: the list creation total run time for this 2<sup>nd</sup> order MLS fit for  $N = N_{max} = 23$  is only approximately 30% longer than that for  $N = N_{min} = 14$ . Figure 42 shows the execution time required by each of the SVD and the GJE solvers versus  $N$ . The total run time for the SVD solver is proportional to  $N^{0.9}$  and the total run time for the GJE solver is proportional to  $N^{0.4}$ .

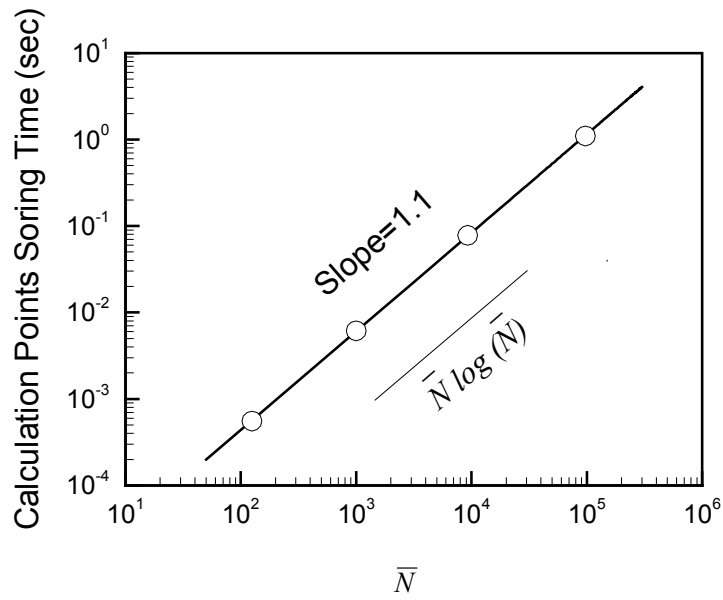


Figure 38. Total time required to sort the calculation points  $\bar{N}$ .

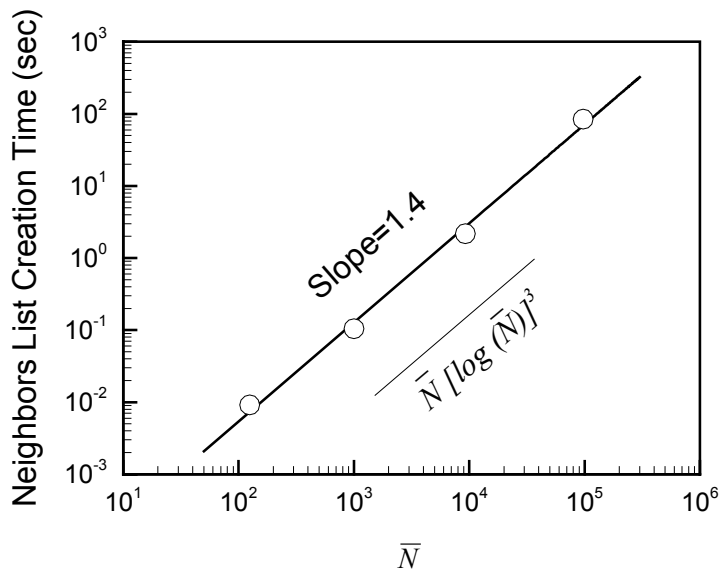


Figure 39. Total time required to create the nearest neighbors lists.

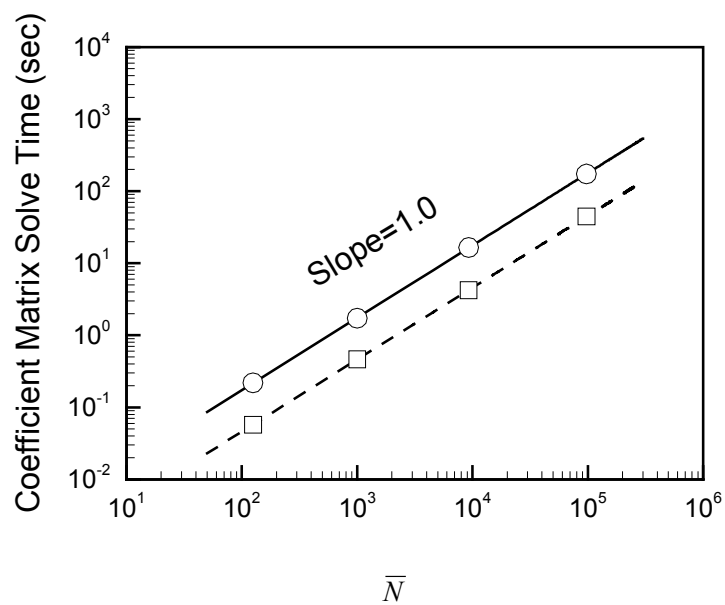


Figure 40. Total time required by the SVD solver (—○—) and by the GJE solver (---□---).

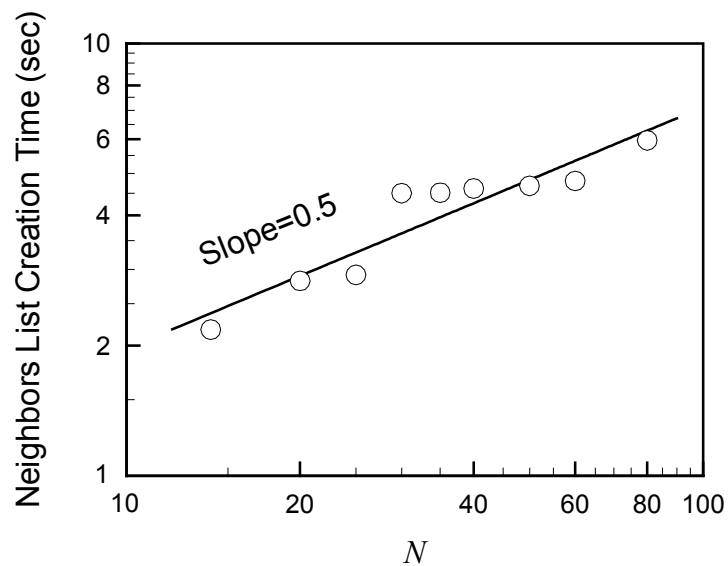


Figure 41. Total time required to create the nearest neighbors lists.

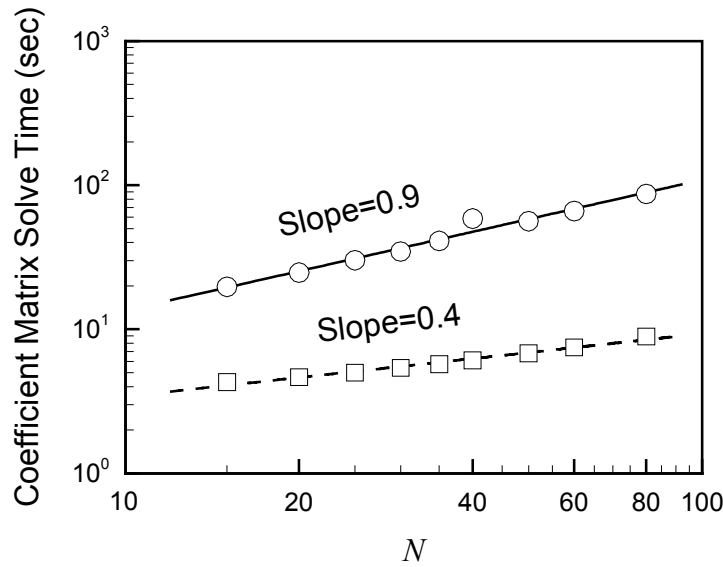


Figure 42. Total time required by the SVD solver (—○—) and by the GJE solver (---□---).

Using the results of this subsection, by way of example, when doubling the number of calculations points in the nearest neighbors used in the MLS fit the total execution time (including the sort, nearest neighbors list and matrix solution) increases by a factor of 2.8 using SVD solver and by a factor of 1.8 using the GJE solver. The total execution time for  $N_s = 21$  and  $N = N_{min} = 14$  is approximately 21 seconds using the SVD solver and approximately 7 seconds using the GJE solver. Increasing the number of nearest neighbors used in the MLS fit to  $N = N_{max} = 23$ , the approximate execution times are, respectively, 32 and 8 seconds.

Ultimately, selection of the number of nearest neighbors  $N$  to be used depends on only two considerations: on whether point spacing is locally randomized everywhere and on the order of MLS fit required. These points are considered in detail in the following section.

### Subsection Summary

(i) The total time required by any linear equation solver used in the MLS approach runs in order  $\bar{N}$  time; that is, in linear time with respect to  $\bar{N}$ . (ii) The total time required to generate the nearest neighbors lists is of the same order as that required by the linear equation solver. (iii) The increased execution time associated with an increased number of nearest neighbors  $N$  is relatively small. (iv) The SVD solver requires approximately 4 times the execution time of the GJE solver.

## 4 SUMMARY

This paper presents an MLS formulation for multi-dimensional applications and presents a detailed investigation of a three-dimensional MLS application including errors in the approximation of 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> order derivatives. The existing technical literature does not include either extensive application analyses or analytical error bounds for MLS approximations of multi-dimensional functions or their derivatives.

Results in this report show that the order of the conserved moment of a functional in a minimization approach such as MLS does not equal the order of accuracy of the approximation. The summary given in Table 4 shows the nominal order of accuracy for an  $h^{\text{th}}$  order MLS fit approximating a  $j^{\text{th}}$  order gradient is roughly  $(h - j + 1)$ . These observed orders of accuracy result from changes in the ratio of average point spacing  $\Delta x$  to a convenient measure of the severity of function curvature such as  $\delta$ .

MLS Order of Fit $h$	$\nabla f$		$\nabla^2 f$		$\nabla^3 f$	
	$E_{rms}$	Order of Accuracy	$E_{rms}$	Order of Accuracy	$E_{rms}$	Order of Accuracy
1	$10^{-2}$	1				
2	$10^{-2} - 10^{-3}$	2	$10^{-2}$	1.5		
3	$10^{-3} - 10^{-4}$	3	$10^{-2} - 10^{-3}$	2	$10^{-1} - 10^{-2}$	1+

Table 4. Summary of the typical  $E_{rms}$  error and the nominal order of accuracy for the MLS application tests in this report.

The SVD solver is required for an MLS fit performed using double-precision calculations only if the number of nearest neighbors is not strictly controllable or if spacing of the calculation points is uniform. For the MLS fits in this report, the use of an SVD solver is never required if the calculation points are even minimally randomized. Furthermore, even *combinations* of extreme conditions considered do not indicate that SVD is required. For example, Table 5 summarizes conditions and results calculated using the GJE and SVD solvers to approximate a severely narrow function (small Gaussian test function radius  $\delta$  corresponding with large wave amplitude to wave width ratio  $\Omega$ ) with low resolution (small

value of  $N_s$  corresponding to a small value of  $\bar{N}$ ), using few nearest neighbors (small value of  $N$ ) with minimum randomization of the calculation point positions (small value of the maximum perturbation  $\varepsilon$ ), and including an externally isolated point (with a value of the scaled isolation distance  $D$  at which GJE fails for uniformly spaced calculations points). As expected, both the condition number and the error in the calculations for such extreme conditions are high; nonetheless, results of the GJE and SVD solvers are identical to 5 decimal places and the GJE solver does not fail.

	Calculation Results			Extreme Parameter Values			
	$C_{max}$	$E_{rms}$ $\nabla f$	$E_{rms}$ $\nabla^2 f$	$\delta / \Omega$	$N_s / \bar{N}$	$\varepsilon$	$D$
GJE	$6 \times 10^4$	1.050989	1.125372	0.05 / 4800	21 / 9261	0.01	4
SVD		1.050990	1.125372				

Table 5. An example of nearly identical calculation results using the GJE and SVD linear equation solvers under extreme conditions.

The question of how many nearest neighbors  $N$  to use can now be addressed with some certainty. If the calculation point spacing is randomized everywhere then only the chosen order of MLS fit determines  $N$ . The requirement of irregularly spaced points precludes the use of interpolation, in whole or part, of the function space onto a regular grid. For example, this rules out an approach to ensure high resolution of a boundary layer by using an Eulerian-Lagrangian scheme in which calculation points are interpolated onto a uniform grid. If the randomness of point spacing is guaranteed, then a direct solver such as GJE can be considered. Using such a solver requires more nearest neighbors  $N$  in the fit to ensure that the solver will never fail as the condition number increases due to redundant equations in the MLS coefficient matrix (particularly at higher orders of MLS fit). However, presented test results show that using additional nearest neighbors, say  $N_{max}$  versus  $N_{min}$  or even  $N > N_{max}$ , is acceptable: there is little sensitivity of  $E_{rms}$  error to  $N$  (including the corresponding changes in the MLS relative window size given by  $\frac{r}{\delta}$ ); for  $N \geq N_{min}$ , an increase in  $N$  shows a reduced sensitivity to randomization of points and an increase in order of accuracy; there is only a small penalty in total run time for an increase in  $N$  yet a direct solver will still execute many times faster than the SVD solver. For example, the direct LU

decomposition linear equation solver executes in approximately 1/3 the time of the GJE solver [15] and so will execute nominally 12 times faster than that SVD solver using the same number of nearest neighbors. Finally, using randomized points, the order of the MLS fit to use depends principally on the order of accuracy required. If a roughly 1.5 order of accuracy is acceptable, a 2<sup>nd</sup> order fit suffices. For any order fit, a conservative number of nearest neighbors to use would then be  $N_{max} \leq N$ .

If the calculation point spacing is not randomized everywhere or if the number of nearest neighbors cannot be specified, the SVD solver may be required. The number of nearest neighbors might not be able to be specified if nearest neighbors are selected from a list (such as a limited connectivity list) generated for other purposes. In such cases, even for randomized calculation point locations, if  $N$  falls below the value of  $N_{min}$  the SVD solver yields results with errors on the order of the peak function value being approximated, though the solver will not fail.

The issue arises whether perhaps the best computational approach is to evaluate the condition number prior to each matrix solution and then to conditionally execute a fast direct solver or the SVD solver. Unfortunately, evaluation of the condition number requires a decomposition of the matrix requiring run time approaching that of the SVD solver.

While condition number  $C_{max}$  and error  $E_{rms}$  values are critically sensitive to *whether* the calculation point positions are uniformly spaced or randomly spaced, these values are not very sensitive to the *degree* of randomization. These values are also not very sensitive to the scaled isolation distance  $D$  of a calculation point. Sensitivity to increasing severity  $\Omega$  of curves produces errors localized to regions of high curvature in the function being approximated such that the approximated curve peaks are reduced in magnitude and the curve widths are broadened.

A broad result is observed for the MLS approach, which may apply to residual minimization approaches in general: compared to calculations using uniformly spaced points, calculations using randomized calculation points can yield both greater accuracy as measured by  $E_{rms}$  and greater stability as measured by  $C_{max}$ .

This page intentionally left blank.



## REFERENCES

1. Savitzky, A., Golay, M. J. E., "Smoothing and Differentiation of Data by Simplified Least Squares Procedures," *Analyt. Chem.*, Vol. 36, pp. 1627-1639, 1964.
2. Ziegler, H., "Properties of Digital Smoothing Polynomial (DISPO) Filters," *App. Spec.*, pp. 1583-1586, 1981.
3. Hamming, R. W., *Digital Filters*, 3<sup>rd</sup> Ed., Chapter 5, Prentice Hall, 1998.
4. Shepard, D., "A Two-Dimensional Interpolation Function for Irregularly-Spaced Data," *Proc. ACM Natl. Conf.*, pp. 517-524, 1968.
5. Lancaster, P., Salkauskas, K., "Surfaces Generated by Moving Least Squares Methods," *Math. of Comp.* Vol. 155, pp. 141-158, 1981.
6. Armentano, M. G., Duran, R. G. "Error Estimates for Moving Least Square Approximations," *App. Numer. Math.*, Vol. 37, pp. 397-416, 2001.
7. Nayroles, B., Touzot, G., Villon, P., "Generalizing the Finite Element Method: Diffuse Approximation and Diffuse Elements," *Comput. Mech.*, Vol. 10, pp. 307-318, 1992.
8. Belytschko, T., Gu, L., Lu, Y. Y., "Fracture and Crack Growth by Element-Free Galerkin Methods," *Modeling Simul. Mater. Sci. Eng.*, Vol. 2, pp. 519-534, 1994.
9. Belytschko, T., Krongauz, Y., Organ, D., Fleming, M., Krysl, P., "Meshless Methods: An Overview and Recent Developments," *Comput. Meth. Appl. Mech. Engrg.*, Vol. 139, pp. 3-47, 1996.
10. Melenk, J. M., Babuska, I., "The Partition of Unity Method: Basic Theory and Applications," *Comp. Meth. Appl. Mech. And Engrg.*, Vol. 139, pp. 289-314, 1996.
11. Duarte, C. A., Oden, J. T., "HP Clouds—An HP Meshless Method," *Numerical Methods for Partial Differential Equations*, pp. 1-34, 1996.
12. Marshall, J. S., Grant, J. R., "A Lagrangian Vorticity Collocation Method for Viscous, Axisymmetric Flows With and Without Swirl," *J. Comput. Phys.*, Vol. 138, pp. 302-330, 1997.
13. Griffiths, D. V., Smith, I. M., *Numerical Methods for Engineers*, CRC Press, p. 262.
14. Zwillinger, D., *Standard Mathematical Tables and Formulae*, CRC Press, p.168, 1996.

15. W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes in Fortran*, 2<sup>nd</sup> Ed., Cambridge Univ. Press, 1992.
16. Miller, R., Boxer, L., *Algorithms Sequential and Parallel*, p. 39, Prentice Hall, 2000.

## DISTRIBUTION

- |  |   |
|--|---|
| 1 Prof. A. J. Chorin<br>Department of Mathematics<br>University of California<br>Berkeley, CA 94720  | 1 Prof. S. Mas-Gallic<br>Centre de Mathematiques Appliquees<br>Ecole Polytechnique<br>91128 Palaiseau Cedex FRANCE                            |
| 1 Dr. Adrin Gharakhani<br>Applied Scientific Research<br>1800 E. Garry Ave., Suite 214<br>Santa Ana, CA 92705                                  | 1 Dr. Jeffery S. Marshall<br>Iowa Institute of Hydraulic Research<br>University of Iowa<br>300 S. Riverside Drive<br>Iowa City, IA 52242-1585 |
| 1 Prof. Ahmed F. Ghoniem<br>Massachusetts Institute of Technology<br>Mechanical Engineering Dept.<br>Room 3-342<br>Cambridge, MA 02139-4307    | 1 Dr. Eckart Meiburg<br>Dept. of Aerospace Eng.<br>University of Southern California<br>854 W. 36th Place<br>Los Angeles, CA 90089-1191       |
| 1 Dr. John R. Grant<br>Naval Undersea Center<br>1176 Howell Street<br>Building 108, Code 8233<br>Newport RI 02841-1708                         | 1 Dr. D. I. Meiron<br>Dept. of Applied Mathematics<br>California Institute of Technology<br>Pasadena, CA 91125                                |
| 1 Prof. Leslie Greengard<br>Courant Institute of Math. Sciences<br>New York University<br>251 Mercer Street<br>New York, NY 10012              | 1 Dr. M. Nitsche<br>Dept. of Math. & Stat.<br>University of New Mexico<br>Albuquerque, NM 87131   |
| 1 Dr. Stephen Huyer<br>Naval Undersea Weapon Center<br>Hydrodynamics Branch, Code 8233<br>Building 1302/1<br>Newport, RI 02841                 | 1 Prof. Y. Ogami<br>Dept. Mech. Eng.<br>Ritsumeikan University<br>Kusatsu 525-77, JAPAN   |
| 1 Prof. Omar M. Knio<br>Dept. of Mech. Eng.<br>Johns Hopkins University<br>Baltimore, MD 21218-2686  | 1 Prof. V. Rokhlin<br>Department of Computer Science<br>Yale University<br>PO Box 2158<br>New Haven, CT 06520                                 |
| 1 Prof. Anthony Leonard<br>Graduate Aeronautics Lab.<br>1200 East California Blvd.<br>California Institute of Technology<br>Pasadena, CA 91125 | 1 Prof. P. G. Saffman<br>Dept. of Applied Mathematics<br>California Institute of Technology<br>Pasadena, CA 91125                             |

- 1 Prof. T. Sarpkaya  
Dept. Mech. Eng.  
Code 69-SL  
Naval Postgraduate Academy  
Monterey, CA 93943
- 1 Prof. J. A. Sethian  
Dept. of Mathematics  
University of California  
Berkeley, CA 94720
- 1 Dr. S. Shankar  
Quantum Corporation  
MS E23  
333 South Street  
Shrewsbury, MA 01545
- 1 Prof. O. R. Tutty  
Dept. of Aero. & Astro.  
University of Southampton  
Highfield, Southampton, SO17 1BJ  
Hampshire, United Kingdom
- 1 Prof. L. Van Dommelen  
FAMU-FSU College of Engineering  
2525 Pottsdamer Street, Room 229  
Florida State University  
Tallahassee, FL 32310-6046
- 1 Prof. G. S. Wincklemans  
Center for Sys. Eng. & App. Mech.  
Unité TERM, place du Levant 2,  
Université Catholique de Louvain  
Louvain-la-Neuve 1348  
BELGIUM
- 1 Prof. Norman J. Zabusky  
Dept. of Mech. and Aerospace Eng.  
Rutgers University  
PO Box 909  
Piscataway, NJ 08855-0909
- 1 0835 J. H. Strickland, 9141
- 1 0835 J. S. Peery, 9142
- 1 0835 M. W. Glass, 9141
- 1 0835 R. J. Cochran, 9141
- 1 0835 S. Burns, 9141
- 1 0835 S. N. Kempka, 9141
- 1 0835 S. R. Subia, 9141
- 1 0835 V. L. Porter, 9142
- 1 0835 W. P. Wolfe, 9115
- 1 0836 C. W. Peterson, 9100
- 1 0836 S. R. Tieszen, 9132
- 1 0836 T. E. Voth, 9116
- 1 0841 T. C. Bickel, 9100
- 1 9018 Central Technical Files, 8945-1
- 2 0899 Technical Library, 9616
- 1 0612 Review & Approval Desk,  
9612 for DOE/OSTI
- 1 0825 W. H. Rutledge, 9115
- 1 0826 W. L. Hermina, 9113
- 1 0827 R. Griffith, 9117
- 1 0828 J. A. Fernandez, 9140
- 1 0828 W. L. Oberkampf, 9133
- 1 0834 A. C. Ratzel, 9110
- 12 0835 A. A. Gossler, 9141
- 1 0835 G. F. Homicz, 9141